SciDAC Visualization and Analytics Center for Enabling Technologies
Project Management Plan
Version 1.2

E. Wes Bethel[*] and Chris Johnson[†]
Principal Investigators

Charles Hansen, Claudio Silva, Steven Parker, Allen Sanderson,
Xavier Tricoche, Martin Cole[‡]

Sean Ahern, George Ostrouchov, Dave Pugmire, Jeremy Meredith[§]

Valerio Pascucci, Hank Childs, Peer-Timo Bremer,
Daniel Laney, Ajith Mascarenhas, Kathleen Bonnell[¶]

Ken Joy, Bernd Hamann[‖]

Cecilia Aragon, Gunther Weber, Prabhat[**]

October 2007

[*] Lawrence Berkeley National Laboratory
[†] Scientific Computing Institute, University of Utah
[‡] Scientific Computing Institute, University of Utah
[§] Oak Ridge National Laboratory
[¶] Lawrence Livermore National Laboratory
[‖] University of California – Davis
[**] Lawrence Berkeley National Laboratory

Contents

1 Mission

The SciDAC Visualization and Analytics Center for Enabling Technologies (VACET) focuses on leveraging scientific visualization and analytics software technology as an enabling technology for increasing scientific productivity and insight. Our mission is to foster scientific insight through creating and deploying effective data understanding technology that is truly responsive to the needs of our stakeholders in the scientific research community. Our stakeholders – researchers from from experimental and computational sciences – are "awash in data" produced by an "information big bang" whereby their ability to collect and generate data far out paces the ability to understand such data. This challenge is widely acknowledged to be one of the primary bottlenecks in contemporary science.

The vision for our Center is to respond directly to this challenge by adapting, extending, creating when necessary and deploying visualization and data understanding technologies for our science stakeholders. Using an organizational model Center for Enabling Technologies, we are well positioned to be responsive to the needs of a diverse set of scientific stakeholders in a coordinated fashion using a range of visualization, mathematics, statistics, computer and computational science and data management technologies.

2 Goals and Objectives

SciDAC's stated mission is to use advanced computing to further the DOE research programs in basic energy sciences, biological and environmental research fusion, and high-energy and nuclear physics. More specifically, the SciDAC program aims to produce "a Scientific Computing Software Infrastructure that bridges the gap between advanced research in applied mathematics and computer science, and computational science research in the physical, chemical, biological and environmental sciences [2]." Stated differently, SciDAC's mission is enable advances in scientific understanding through use of high performance computational platforms.

SciDAC will achieve its mission through several interrelated thrusts. First, Science Applications produce computational codes that simulate scientific phenomenon ranging from models of exploding supernovae to predicting function of new protein structures. Second, Centers for Enabling Technology (CETs) produce computational infrastructure that play varying roles in helping Science Applications to achieve their goals. Projects here include core numerical solvers suitable for use on large parallel machines, visual data analysis technology to help Science Applications derive understanding from data produced by their codes to performance models that predict how well a given software application will perform on a given architecture. Third, Partnerships are focused efforts involving personnel from Science Applications, Centers and possibly Institutes (below). These are teams of experts from different areas that combine forces to produce novel solutions. Fourth, Institutes (according to the FY06 call[1] are somewhat more broadly defined: University-lead centers of excellence intended to complement CET's that [...] focus on major software issues through a range of collaborative research interactions. Stated differently, Institutes serve as a focal point for research activities in a given area, reach out to engage of broader activities of scientific discovery, have a dimension of training and outreach in high performance computing topics including graduate students and postdocs.

The SciDAC program as a whole is evaluated by its quality of scientific output. In the case of Science Applications, this means new scientific discoveries. More broadly, it means all elements of

[1] http://www.science.doe.gov/grants/FAPN06-04.html

the SciDAC program are expected to produce top quality research as evidenced by peer-reviewed publications.

As a CET, one of VACET's central goals is to bring cutting-edge, production quality visualization and analytics to bear on the data understanding challenges of the SciDAC scientific community in the form of well engineered, usable software. Achieving this overall mission objective will entail accomplishing several interrelated goals. Achieving these goals is possible only through our organization as a SciDAC Center for Enabling Technologies, as described later in Section 3.

2.1 Foster Effective Stakeholder Relationships

One of the central pillars of our Center is strong relationships with our science stakeholders. The specific goal here is achieving vibrant, productive interactions with our customers/stakeholders. We translate their needs into specific, actionable items that in turn result in new scientific data understanding capabilities.

2.2 Coordinated Interactions with Stakeholders

Like most within the scientific research community, our stakeholders' time is scarce and valuable. Truly effective data understanding solutions emerge only through meaningful and ongoing dialog between stakeholder and technology provider. Our stakeholders typically do not have the luxury of duplicating such interactions across a number of different data understanding providers. One goal of our center, which is possible only through organization as a Center for Enabling Technologies, is to focus stakeholder interactions through a VACET "Customer Project Manager," who is responsible for such interactions between stakeholder and VACET. We have organized our center as described in Section 6, to optimize the all-important stakeholder-VACET relationship, which is a crucial element of success to both stakeholder and VACET.

After our first year of operation, this organizational model has proven to be very successful: not only is our Center very productive, our stakeholders have been vociferous in their support and approval of this type of organization.

2.3 Effective Production Deployment of Petascale-Capable Technology

One of VACET's goals is to deploy its production-quality, petascale-capable data understanding technology at DOE's Open Computing facilities. To that end, VACET team members include team leads from visualization and analytics at the National Energy Research Scientific Computing Center (NERSC) at Lawrence Berkeley National Laboratory and the Leadership Computing Facility (LCF) at Oak Ridge National Laboratory. A motivation for such a focus is the fact that most of DOE's computational science activities are conducted at these facilities. We anticipate in the future an increasing trend towards consolidation of community-wide data repositories for computational and experimental sciences at these facilities. Therefore, we have engineered into our team's organization and makeup close relationships with production visualization and analytics activities at these DOE computing facilities.

2.4 Coordinated Interactions with Other SciDAC Technology Providers

One of the strengths of the SciDAC model is the opportunity for leveraging multiple technologies to provide focused solutions. In this spirit, VACET aims to leverage as much existing technology from other SciDAC technology providers as appropriate to provide integrated solutions to our stakeholders by reducing our development time and accelerating our time to solution. Activities

towards achieving this goal strengthen not only VACET but other individual SciDAC efforts as well as the overall SciDAC program.

2.5 Technical Excellence, Publications, Outreach, Education and Service

We anticipate that most of VACET's accomplishments will be groundbreaking, "first-ever" works that span a diverse range of research, development, engineering, deployment, and novel application to challenging data understanding problems. As a consequence, our team feels it important to achieve a high level technical excellence and correspondingly a leadership role in the field of high performance visual data understanding. As communication of results is an important element of our work, we expect our team members to be high achievers in terms of technical publications to convey our accomplishments and the success of the SciDAC program.

Outreach and education activities help to increase the "surface area" between SciDAC, VACET and the scientific community as well as the general public. Outreach activities include things like invited talks, workshops and tutorials. Education activities include opportunities for students to engage in work on "real world" problems, to be exposed to a highly productive, diverse and dynamic team comprised of leaders in the field of visual data understanding. Service includes activities like conference chair, program committee member, and technical paper/proposal reviewer for journals, conferences and DOE programs.

2.6 Train the Next Generation of Visualization and Analytics Leaders

Related, we feel an important VACET goal is to train the next generation of leaders in the field of high performance visual data understanding. In addition to the opportunities for young student scientists, our Center is organized (see Section 6) in such a way as to provide a great deal of opportunity for interaction with stakeholders, other VACET team members, and, perhaps most importantly, opportunity for leadership roles within the Center in the form of Customer Project Manager and technical project Team Leader. All VACET team members participate in project management activities, and all are expected to generate technical publications (papers, web pages, etc.) and to prepare and deliver presentations.

2.7 Provide Positive Return on SciDAC Investment

Since this effort is the first major coordinated visualization activity in DOE in recent years, all on our team are committed to ensure that DOE's investment pays handsome dividends in the form of positive and significant scientific impact.

3 Benefits

SciDAC's stated mission is to use advanced computing to further the DOE research programs in basic energy sciences, biological and environmental research fusion, and high-energy and nuclear physics. More specifically, the SciDAC program aims to produce "a Scientific Computing Software Infrastructure that bridges the gap between advanced research in applied mathematics and computer science, and computational science research in the physical, chemical, biological and environmental sciences." [2].

The first five years of the SciDAC program (2001-2006) have produced dramatic advances in scientific software technology that can take advantage of the world's most powerful computing systems. While most of the SciDAC investment portfolio has focused on scientific software and

computing, a small fraction has been dedicated to management of the resulting data. In the first round of SciDAC, there was no organized or coordinated data understanding or visualization effort. The result is a scientific community that has the ability to generate vast amounts of data and some ability to store and retrieve data, but only a minimally coordinated, effective, deployed, and supported set of software tools and techniques for data understanding.

While the general benefits of scientific visualization are generally well understood and accepted, there are a number of unique benefits that result from our team organized and operated as an integrated Center for Enabling Technologies within the SciDAC program.

3.1 Integrated, Coordinated Visualization and Analytics R&D Activities

Visualization and analytics activities within DOE have traditionally been conducted as small, uncoordinated efforts. With a very few isolated exceptions, most notably the concerted visualization effort within the ASC program [1], there has been no impetus for a DOE-funded visualization scientist to go beyond publishing a research paper, making a pretty picture, and occasionally producing a stand-alone application. The approach we take here is to consolidate and coordinate efforts of leading visualization researchers to support the open science community. Our mission involves the coordinated activities of several different groups and thus requires long term funding to enable us to foster and maintain relationships with application scientists. Such relationships ensure that visualization solutions are responsive to stakeholders' needs, are deployed in the appropriate places, and are supported over the course of a given science project. This type of approach, and the resulting benefit, is possible only through organization as a Center; such results are not possible at the scale of the individual project.

3.2 Leveraging Other SciDAC Technologies to Accelerate Progress

As a Center, we can coordinate the development and deployment activities of related and complementary technologies such as data management, analysis, and visualization. Further, as technologies are developed for a specific application, we will be in a position to rapidly capitalize upon the opportunity for broader deployment in other application areas. This type of coordinated effort is not possible at the small project scale of the individual PI. As a Center, we can serve a greater number of stakeholders than would be possible if organized as a number of smaller, isolated efforts.

3.3 Integrated and Effective Deployment

As a Center, we are in a position to interact with other technology centers – Mathematics and Computer Science – to ensure that our data understanding technology is deployed as part of a larger technology ecosystem. Again, that level of broad interaction is not possible at the level of the individual PI. Related, our team includes leaders from production visualization and analytics efforts at NERSC/LBNL and LCF/ORNL, thereby maximizing opportunity for successful deployment at those facilities, which are host to many SciDAC and other computational and experimental sciences.

3.4 Effective Solutions from a Broad Technology Base

Our experience has shown that there simply is no "single visualization technology solution" (no "one size fits all" solution) that is responsive to the broad set of challenges facing the scientific research community. Instead, effective solutions require the careful adaptation and deployment of technologies from many sources. Some technologies, such as application frameworks, are general

purpose, while other technologies, like visual presentation paradigms showing the metabolic pathways within a bacterium, are uniquely specific to a given application area. The ability to vet a broad technology base within a Center helps ensure effective solutions and also protects DOE's investment in those constituent technologies.

3.5 Enable Scientific Discovery

With the creation, blending, application and deployment of technology, we are in a unique position to provide markedly new scientific data understanding capabilities. Our operational model, which relies on substantive and long-term relationships with stakeholder/customers, ensures our work focuses on addressing scientific data understanding needs, is engineered to deliver scientific impact.

3.6 Time and Cost Savings

In the past, some of our customers have been forced to allocate a portion of their scarce science research budget to create visualization and analytics infrastructure. Unanimously, these stakeholders desire to get out of the business of developing and maintaining visualization infrastructure. One aspect of our strategy is to provide community-wide visualization infrastructure that will thereby result in cost and time savings to scientific applications: time and effort previously spent on building and maintaining visualization infrastructure can instead be expended on science rather than software engineering.

We expect to realize time and cost savings in many other areas as well. For example, our organization as a Center allows us to focus R&D efforts into core technologies that will then be applied to a variety of different customer projects. A long funding horizon allows us to exercise technology options over different time scales from short- to long-term. Such a model – centralized and coordinated R&D that is then applied in "products" reflects successful business models in industry.

4 Problem

Since the advent of computing, the world has experienced an "information big bang", an explosion of data. Information is being created at an exponential rate, such that since 2003, new information generated annually exceeds the information contained in all previously created documents. Furthermore, digital information now makes up more than 90% of all information produced, vastly exceeding that generated on paper and film. One of the greatest scientific and engineering challenges of the 21st century is to effectively understand and make use of this growing wealth of information to scientific advantage. Software plays a central role in providing the means to manage and understand relationships, anomalies, trends, and features contained in today's abundant scientific data.

In the past decade, the art and science of data understanding in general, and scientific visualization in particular, have been subject to the same challenges as all other scientific fields affected by the explosive growth in the size and complexity of data. The visualization community has responded with a diverse range of research spanning the creation of scalable techniques for increasing visualization throughput, the development of techniques for multivariate visual display and exploration, and a host of other activities [34, 25, 52, 32]. Unfortunately, many of these activities have been conducted in a vacuum, without coordination, with little or no input from stakeholders, and with little emphasis upon software deployment or on a given technique's actual effectiveness in terms of scientific understanding [26]. There have been many examples of "pretty pictures"

that serve well as vehicles of communication, but such visualization results often do not necessarily produce substantive forward progress in terms of scientific understanding.

Three primary challenges face the SciDAC visualization and data understanding communities in their quest to have a substantive impact on science and engineering. First, visualization and analytics research and development must be more attuned to the needs of SciDAC scientists and engineers. While creating scalable visualization tools is an important part of a broad strategy for meeting those needs, simply creating an image of a terabyte's worth of raw scientific data does not guarantee increased scientific insight. Indeed, if the resulting image is too complex, it may even reduce the likelihood of discovery.

Second, visualization and analytics technology must be designed and implemented as part of a larger technology ecosystem in which it is integrated with other important data management and work-flow systems to better meet the data understanding needs of the scientific community. They need to be integral with data acquisition, management, storage, and retrieval, and be a part of a researchers work-flows and systems rather as separate component parts or isolated islands of capability. Visualization and analytics technologies that are engineered to leverage complementary technologies in data management, parallel computing, and software engineering are much more likely to succeed at their intended purpose and are therefore more likely to be used by scientists as part of their day-to-day investigatory methodologies.

Third, visualization and analytics technology must be deployed, maintained, and supported in the computing environment. It is not enough simply to publish a visualization research paper and post source code on a web site. Scientific researchers need expert help from visualization and analytics scientists to deploy, tune, and adapt technology for their specific scientific domains. VACET's mission and scope directly addresses all of these important challenges.

5 Approach

Our proposed work plan directly responds to the data understanding needs of our scientific stakeholders and will produce results that meet those needs. As a Center, we are well-positioned to coordinate development and deployment activities across a spectrum of data understanding technologies and groups as well as with other SciDAC centers. This approach ensures our team's results are part of a well-balanced program within a larger technology ecosystem.

5.1 Provide Production-Quality, Petascale-Capable Visual Analysis Software Infrastructure

SciDAC's stated mission is to use advanced computing to further the DOE research programs in basic energy sciences, biological and environmental research fusion, and high-energy and nuclear physics. More specifically, the SciDAC program aims to produce "a Scientific Computing Software Infrastructure that bridges the gap between advanced research in applied mathematics and computer science, and computational science research in the physical, chemical, biological and environmental sciences." [2].

5.2 Visual Data Analysis as an Integral Part of the Scientific Process

Scientific visualization plays an important role in the scientific process: as "virtual oscilloscope" to "see the unsee-able"; as the most visible element of scientific research and insight; and as the visual component of day-to-day diagnostic and exploration tools. Its aim is to help scientists gain

insight into structures, relationships, and anomalies "hidden" within data. Its impact can be seen in all areas of science, medicine, engineering, finance, security, and safety.

5.3 The Customer/Stakeholder Relationship

Our team has strong relationships with application scientists. Additionally, some of our team members are shared with DOE's large-scale open computing facilities (National Energy Research Scientific Computing Center (NERSC) at Lawrence Berkeley National Laboratory and the Leadership Computing Facility (LCF) at Oak Ridge National Laboratory), which helps to maximize the likelihood that VACET's new technologies will be successfully deployed at those centers for stakeholder use. Most members of our team have been active in conducting visualization research, development, and deployment activities during the first five years of SciDAC. We are collectively dedicated to our Center's mission: To foster scientific insight through creating and deploying effective data understanding technology that is truly responsive to the needs of the scientific research community.

The Customer/Stakeholder relationship comprises one of the foundational elements of our overall approach. These relationships, are long-term, collaborative and mutually beneficial to both stakeholder and VACET. Most of the work we perform in VACET is aimed at meeting scientific data understanding needs. Those needs are articulated by the stakeholder to use within the context of these relationships. The solutions are then tested and applied on stakeholder problems to meet scientific data understanding needs, and where appropriate, included in releases of production quality software.

5.4 Deployment Approaches

In the way there is no "one-size-fits-all" solution to scientific data understanding, there is similarly no one deployment solution that will satisfy all stakeholder operational requirements. We adopt a set of tactics, described in the following section, that affords us the framework needed to achieve maximum impact, maximum mobility and flexibility in delivering solutions, and optimizes the pipeline from concept to working software solution.

5.4.1 Applications

One of the primary strengths of our team's approach is evidenced by the fact that we are delivering production-quality solutions to customers after less than six months since VACET began business. The approach we use to accomplish this remarkable achievement is by leveraging a large body of production-quality software infrastructure as the base for both R&D and delivering new capabilities. Our two primary delivery platforms consist of VisIt and SCIRun.

The SCIRun system has been a focus of research and development at the SCI Institute since 1995 [28, 42, 44, 41, 43, 27, 37, 36, 29]. It is framework for visualization, modeling, and simulation, and has been the test bed for significant fundamental research in visualization techniques and their applications to real-world scientific problems. The strengths of SCIRun derive from its modular data flow architecture, which provides a much wider range of flexibility via modular pipelines and *dynamic compilation*. SCIRun2 [45] expands dramatically on these ideas to bring component-based scientific computing and visualization to an entirely new level. The primary novel feature in SCIRun2 is the concept of a metacomponent, [60] which allows construction of scientific software that involves mixtures of components from different sources, including support for the Common Component Architecture (CCA) [5], the Visualization Toolkit [50], CORBA [39], and dataflow components from the original SCIRun. Components from these different sources can be combined in

a single computation via the use of automatically- or semi-automatically-created bridges. SCIRun2 also enables parallel components through multi-threading for shared memory programming, and parallel-to-parallel remote method invocation [10, 19] for connecting components in a distributed memory environment.

VisIt [17, 18] is a turnkey application for data exploration, code assessment, and quantitative analysis suitable for use on tera- and peta-scale datasets. In addition to standard visualization methods, the tool is used for code-to-code comparisons, code-to-experiment comparisons, analysis of parameter studies, and quantitative analysis across a variety of scientific areas. It has won an R&D 100 award, been downloaded over 25,000 times, has over 300 customers at LLNL, is used at AWE, many ASC Alliance sites, and has a large number of customers from businesses, universities, and national laboratories. Originating in 2000 as part of the ASC program, it has grown to over one million lines of code in addition to leveraging many third-party libraries. Its development has focused on key areas where solutions do not already exist: large data infrastructure, unusual data models, custom and extensible quantitative analysis and the infrastructure that binds them together. It has a scalable architecture for running expensive (I/O or compute) operations on a parallel machine to leverage resources "close to" the data and supports a client-server model for effective remote visualization use. It is extensible – developers can write plug-ins for any stage of I/O, visualization or analysis processing. Such plug-ins may run in parallel, thus providing a stable development environment for new techniques in scalable visualization and analysis.

Both tools perform standard visualization operations: contouring, volume rendering, slicing, etc., in addition to extensive subsetting, derived quantity, and analysis capabilities. VisIt has been tailored for processing large data and its server has been parallelized for distributed memory environments. As a result of this expertise and experience, our team is well positioned to respond to our science stakeholder needs.

5.4.2 Libraries and Modules/Components

We also recognize the different needs of the users in terms of delivery vehicles of new technology. Some have already committed to a particular visualization tool and need new capabilities like high performance volume rendering or feature extraction. Others require new focused tools that address specific needs such as real time monitoring of large scale simulations or remote data access and collaborative data exploration. Many users can be well served with general purpose visualization tools, such as VisIt or SCIRun, and we'll provide extensions to support new features and special data formats. We will address this diversity of needs with a three-tiered strategy with development of: **libraries** with narrowly scoped capabilities (e.g. a volume rendering library) and simple APIs that allow easy integration even if at some cost in space and/or time efficiency; **components** that can be easily combined in a data flow network exploiting the efficiency of the parts at their best; and **fully featured application visualization tools** with a large set of features that can satisfy a wide range of users.

This deployment strategy, together with extensive documentation, examples, and tutorials, will also facilitate the process of dissemination to a larger community independent of our ability to provide direct support to each of them.

5.5 Technical Portfolio

The strength of our overall technical approach is based upon a diverse portfolio of capabilities. Since there is no one panacea to the challenges of petascale data understanding, we are leveraging a set of investments across a diverse range of interrelated focus areas. These areas combine in

various combinations to produce a final solution for a given stakeholder. Again, our organization as a Center affords us the opportunity to manage such a portfolio to deliver effective solutions across a wide range of science disciplines, each with slightly different needs.

The details of our technical portfolio exceed the scope of this document. In general terms, the focus areas of our technical portfolio include:

Advanced Visualization Techniques

- Integration of Fundamental Tools for Visualization and Analysis
- Publication Quality Images, Illustrations and Movies
- Project-Wide Visualization Tools
- Flow Visualization
- Scalable Visualization, Analytics and Rendering Tools
- Remote Data Access and Streaming Techniques
- Collaborative Tools

Analytics and Knowledge Discovery

- Query-Driven Visualization
- Feature Detection and Identification
- Integration of Statistical Information
- Information and Scientific Visualization Techniques
- Multifield Visualization and Analysis
- Comparative Visualization and Analytics
- Uncertainty Visualization

6 Team Organization

As described in the VACET proposal, our Center is organized into functional groups to achieve several distinct objectives: (1) facilitate the flow of information between the Center's leadership, personnel, and science stakeholders; (2) to provide the organizational structure needed to ensure oversight and coordinated operations of the Center's collection of activities; (3) to ensure we meet our work deliverables; (4) to gracefully accommodate future growth and respond to changing priorities. The Center's original functional groups are: the Center PIs, the Executive Committee, the External Advisory Board, Research and Development, Chief Software Engineer, Software Development and Support, and Stakeholder Projects. Due to a 45% reduction in budget, we have eliminated the formal External Advisory Board.

Figure 1 depicts the working interaction between these functional groups. The co-Directors (Bethel and Johnson) are responsible for the operation of the entire Center and are the point of contact between the Center and DOE. The Executive Committee (EC), which makes decisions about research, development, stakeholder applications selection and prioritization, and software deployment, is comprised of the participating Laboratory or University PIs (Bethel, Johnson, Pascucci, Ahern, and Joy).

Given the high priority on delivering useful software to our scientific stakeholders, each lab site will have a primary software engineer whose duty it is to assure software developed and deployed from that site meets the Center software engineering criteria. The set of site engineers is known collectively as the Software Engineering Group (SEG). In addition to developing, testing, documenting and maintaining the Center software, the Software Engineering Group will integrate

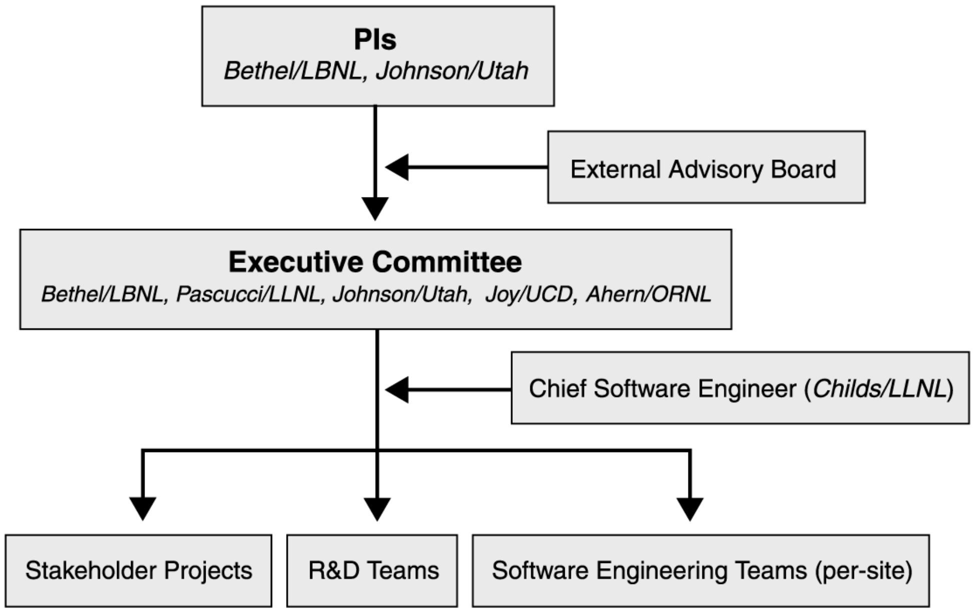

Figure 1: Functional interaction of groups within VACET.

results from all research and development groups into the Center software, relying on feedback from the project leads in the Stakeholder Projects Group, the Executive Committee and the Chief Software Engineer (Childs).

The Chief Software Engineer (CSWE) serves many important functions within the center. One is to facilitate the coordinated design, implementation and integration of the Center's technologies into software solutions that meet stakeholder needs. He will provide guidance to the software development teams so that individual software tools and libraries will be readily usable throughout the Center's collection of stakeholder projects. He will coordinate with the EC to prioritize software development targets, and serve as a technical software advisor to the Center as a whole. He will interact with the R&D team to help foster early designs that fit well within the Center's technology implementations. He will direct the development, testing, deployment, and support of the Center software toolsets.

The Stakeholder Projects Group (SPG) is the primary interface to our science stakeholders. In this group, individuals from the Center will interact directly with science stakeholders to obtain and prioritize science needs, coordinate with the Center's EC and Chief Software Engineer to translate those needs into a work plan, to oversee and manage the work so that software is delivered to the science stakeholder.

From each participating institution is a "Site PI." That person is nominally part of the EC, and is responsible for managing issues at that organization. These issues range from responding to administrative requests, like updated budgets, as well as for managing operations of VACET staff at that organization.

While our original proposal called for a team organization that includes an External Advisory Board (EAB), our project was funded at about 55% of the original request. Therefore, we have eliminated the formal EAB within the organization. Instead, we rely on direct feedback from stakeholders. After our first year of operation, this approach seems to be one that is viable, though not optimal.

7 Roles and Responsibilities

7.1 LBNL

7.1.1 Bethel

As coordinating PI for VACET, Bethel's duties include overall management of a five-institution project, direct management and supervision of LBNL staff contributing to the effort, and interfacing with DOE on all matters relating to the project. In addition to project management duties, Bethel will serve as a Stakeholder Project leader, translate Stakeholder needs into project deliverables, prioritize work tasks, design and implement visualization and analytics solutions particularly in the areas of remote, distributed and parallel visualization, query-driven visualization and analysis, and latency-tolerant visualization and rendering.

7.1.2 Aragon

Dr. Cecilia Aragon is the Team Leader for the Spectrum Synthesis Visual Data analysis project for our Astrophysics stakeholders. Aragon is also Team Leader for the Nearby Supernova Factory software infrastructure, which includes data management, visualization and visual analytics. She is shared with the NERSC Analytics team, the Nearby Supernova Factory and the LBNL Visualization research effort.

7.1.3 Siegerist

Cristina Siegerist left LBNL as of June 2007, and is no longer part of the team.

Her role was to provide one-on-one consulting in the application of visual data analysis technologies to stakeholder projects. She is shared with the NERSC Analytics team.

7.1.4 Prabhat

Prabhat joins the LBNL team as of September 2007. He will assume some of Stockinger's duties in the nexus between scientific data management and visual data exploration. He will also act as an individual contributor to stakeholder projects that are shared with the NERSC Analytics program.

7.1.5 Stockinger

Dr. Kurt Stockinger left LBNL as of September 2007, and is no longer part of the team.

His background was in Scientific Data Management, and his role was to serve as a bridge between efforts in visualization, analytics and data management.

7.1.6 Ushizima

Dr. Daniela Ushizima joins the LBNL team as of October 2007. She is shared with the LBNL Mathematics Department, where she works on projects in the areas of computer vision, machine learning and data analysis. She will contribute to analysis expertise to a number of different VACET projects, starting with astrophysics spectra analysis.

7.1.7 Weber

Dr. Gunther Weber is team-leader-in-training for the Production AMR Visualization effort in VACET. He is also the Team Leader for the 3D LIC R&D effort. He is shared with the NERSC Analytics Team and the LBNL Visualization research efforts.

7.2 ORNL

7.2.1 Ahern

Sean Ahern is the VACET PI for Oak Ridge National Laboratory (ORNL). In addition to directly managing the contributions of the personnel at Oak Ridge to VACET, Ahern serves as a direct customer liaison to users sited at ORNL, as well as users of ORNL's Leadership Computing Facility (LCF). In this role, Ahern translates Stakeholder needs into VACET deliverables, prioritizes ORNL work tasks, and designs and implements visualization and analytics solutions, particularly where they impact users of the LCF. Specific ORNL customers are Dr. Don Batchelor in fusion simulation and Dr. John Drake in climate simulation.

7.2.2 Ostrouchov

Dr. George Ostrouchov is a statistician working for the Computer Science and Mathematics division of ORNL. In his role in VACET, Ostrouchov serves as a liaison to customers in the area of statistical analysis, feature detection, and coupled statistics/visualization deployment. He directly liaises with users located at ORNL. Dr. Ostrouchov also has continuing collaborations with the Scientific Data Management Center for Enabling Technology, allowing for cross-fertilization between the two centers.

7.2.3 Pugmire

Dr. David Pugmire is a computer scientist working for the Center for Computational Sciences at ORNL. Joining VACET in September 2007, Pugmire liaises directly with VACET and ORNL customers. He is a primary VisIt developer especially in the area of custom interfaces and distributed data visualization.

7.2.4 Meredith

Jeremy Meredith is a computer scientist working for the Computer Science and Mathematics division of ORNL. He has extensive experience working with distributed parallel visualization and data analysis codes, especially the VisIt visualization system. He serves as one of the primary code developers for the VisIt system, integrating new analysis capabilities, and customizing data readers for various VACET customers.

7.3 LLNL

7.3.1 Pascucci

Dr. Valerio Pascucci is the coordinator for VACET activities at Lawrence Livermore National Laboratory (LLNL) and will have responsibility for direct management and supervision of LLNL scientists involved in the project. In this capacity, Pascucci will translate and prioritize application

stakeholder needs into LLNL project deliverables, design and implement data analysis and visualization solutions particularly in the areas of robust topological analysis, feature extraction and tracking, linked views, remote, and distributed streaming techniques.

Pascucci will also serve as a Technical Point of Contact for the climate modeling application area and will work with Dean Williams to identify VACET deliverables to be enacted in collaboration with the Earth Science Grid and with John Drake to identify VACET technology that can directly support climate modeling science applications.

7.3.2 Childs

Dr. Hank Childs is the VACET Chief Software Architect and is responsible for the implementation of the activities defined by the Executive Committee with respect to software deployment and distribution as well as quality control.

Childs is also the main LLNL developer responsible for introducing new features in VisIt, such as full support AMR data to satisfy the needs of the APDEC center and relative science applications.

7.3.3 Bremer

Dr. Peer-Timo Bremer is a Computer Scientist at LLNL. He will be the main person responsible for developing robust topological techniques for data analysis and feature extraction for structured and unstructured data. Bremer will work with scientists in different applications to develop formal topological characterization of features of interest and corresponding implementation of discrete algorithms for their computation while providing guaranteed error bounds.

7.3.4 Laney

Dr. Daniel Laney is a Computer Scientist at LLNL. He will be the main person responsible for management of large scientific datasets with particular focus on multiresolution streaming techniques. In particular, Laney will work on the use of progressive techniques for the analysis and visualization of large models from petascale science applications.

7.3.5 Mascarenhas

Dr. Ajith Mascarenhas is a Postdoc at LLNL. He will be the main person responsible for developing robust topological techniques for time dependent datasets and will work on reliable feature tracking methods. His work will include advanced techniques enabling the use of a multi-scale notion of time and allow performing direct selection and filtering of features of in a space-time continuum.

7.3.6 Bonnell

Kathleen Bonnell is a Computer Scientist at LLNL who has extensive expertise developing scientific visualization techniques and who has a deep understanding of the VisIt visualization platform. Bonnell will be responsible for integrating a number of new visualization components and data readers within VisIt as well separating particular VisIt portions into stand-alone libraries that can be deployed separately in other environments.

7.4 Utah

7.4.1 Johnson

Professor Chris Johnson is the VACET Co-PI. Along with VACET Co-PI Wes Bethel, he will lead the overall VACET effort. In addition, Johnson will help create new visualization techniques for displaying error and uncertainty, as well as new multi-field visualization techniques.

7.4.2 Hansen

Professor Hansen and a graduate student will be working on multi-field visualization and uncertainty visualization, focusing on GPU methods for volume rendering for multi-fields taking uncertainty into consideration.

7.4.3 Silva

Professor Claudio Silva will pursue the deployment of VisTrails in the climate application for the purpose of comparative analysis. He will be working with Valerio Pascucci at LLNL, who is the direct contact with Dean Williams (climate stakeholder).

7.4.4 Parker

Professor Steven Parker will work on adapting Utah's real-time ray tracer to particle-based datasets (stakeholders in fusion, astrophysics, accelerator modeling). He will also be working with the CPES center to implement a web-based visualization tool to automate routine visualization tasks for the two-dimensional XGC simulation code. Using Web 2.0 technologies, we will enable viewing large numbers of particles and mesh points in a normal web browser.

7.4.5 Sanderson

Dr. Allen Sanderson's role in VACET will be two-fold: (1) a liaison between VACET and fusion and astrophysics scientists addressing their needs as a stakeholder; (2) developing and deploying visualization and analysis tools based primarily on the needs of fusion and astrophysics stakeholders. During the first year, Sanderson will focus primarily on deploying GPU and volume rendering techniques for visualizing and analyzing particle-based datasets.

7.4.6 Tricoche

Dr. Xavier Tricoche relocated from Utah to Purdue in the Summer of 2007, and is no longer officially part of the VACET team. His role included contributions in the areas of multi-field and flow visualization algorithms.

7.4.7 Cole

Martin Cole is a staff software engineer who will work on transitioning visualization research algorithms to end user software.

7.5 UC Davis

7.5.1 Joy

Professor Joy's duties involve the interaction with stakeholders to identify fundamental visualization research and development problems, to design and translate these problems into workable solutions, and to develop deliverables that can be returned to our stakeholders to address their data analysis needs. His initial efforts will focus on highly-parallel methods for material interface reconstruction, query-driven visualization, and multi-dimensional visualization. These efforts are targeted at stakeholders in combustion, astrophysics, and applied mathematics.

7.5.2 Hamann

Hamann's duties involve investigating topological solutions and multiresolution methods to support analysis and exploration of petascale data sets. He will focus on out-of-core methods and visual segmentation methods, which allows one to represent massive data sets by a relatively smaller number of elements, and implementing these techniques within the frameworks supported by VACET. Through these efforts, he will collaborate with VACET's LLNL component (Pascucci), to adapt and deploy these methods to our VACET stakeholders.

7.5.3 Garth

Dr. Garth joined the UCD team in the Summer of 2007. His research interests include computer graphics, visualization systems, flow visualization, and feature analysis in flow fields.

8 Interrelationships with Other Projects – Centers, Institutes, Applications

8.1 Science Applications

8.1.1 Accelerator

- R. Ryne, LBNL. *Advanced Computing for 21st Century Accelerator Science and Technology.* VACET team members contribute to a Science Application Partnership project that focuses on a community-wide data I/O library. The overall objective is to foster a community data standard so that researchers can share data between teams, as well as between different codes in an ensemble model. An additional benefit is reduced effort for applying advanced visual data understanding technology to this science area.
- P. Spentzouris, FNAL. *Community Petascale Project for Accelerator Science and Simulation (COMPASS)* This accelerator project, which was recently awarded funding by SciDAC, represents a major future stakeholder thrust area for VACET. We have engaged in several early stakeholder needs assessment meetings with several principals from this team. Many of their needs overlap with other those of other stakeholders, notably, those projects in Fusion that generate particle-based data (see Sections 12.1.1, 12.5.2, 12.9.3, 12.5.1, 12.7.5, and 12.9.2).

8.1.2 Astrophysics

- S. Woosley, UC Santa Cruz. *Computational Astrophysics Consortium: Supernovae, Gamma Ray Bursts, and Nucleosynthesis.* VACET is providing production quality, AMR visual data analysis infrastructure (Section 12.6.1) as well as performing R&D for new visual analytics techniques appropriate for Spectrum Synthesis analysis (Section 12.2.1).

8.1.3 Climate

- J. Drake, ORNL. *A Scalable and Extensible Earth System Model for Climate Change Science.* Drake is the consumer of new technologies for climate data analysis and visualization.
- D. Randall, Colorado State. *Design and Testing of a Global Cloud-Resolving Model.* Randall and his team are a prospective VACET collaborator. A member of Randall's team attended the VACET AHM in Salt Lake City in February, 2007.

8.1.4 Combustion

- John Bell and Marc Day, LBNL/CCSE. Bell, a member of the National Academies of Science, performs basic combustion research under DOE base program funding in Mathematics. He is a customer of Colella's Applied Partial Differential Equations Center. Our work with Bell focuses on developing and deploying production quality AMR visual data analysis infrastructure (Section 12.6.1) as well as advanced visual analysis capabilities aimed at gaining traction on difficult scientific data understanding problems (Sections 12.4.1 and 12.4.3).
- Jacqueline Chen and Evatt Hawkes, SNL-CA. *Direct Numerical Simulation of Turbulent Non-Premixed Combustion – Fundamental Insight Towards Predictive Modeling.* VACET's interactions focus on R&D of techniques for topologically based feature identification, tracking and visualization. Traditional visualization methods to not provide the kind of hard, quantitative information needed to advance combustion science. Chen's work is conducted under DOE's INCITE program.

8.1.5 Fusion

- S. Ethier and W. Lee, PPPL. *Center for Gyrokinetic Particle Simulations of Turbulent Transport in Burning Plasmas.* In this active project, VACET is providing novel visual data analysis tools for examining orbitals in very large particle-based datasets. In new work, which is the subject of a May 2007 SAP proposal, we would focus on applying query-driven visualization techniques to analysis of large MHD datasets, and including visual data analysis techniques as part of scientific workflow (this work would be joint with the SDM Center).
- S. Kruger, Tech-X. *Center for Extended Magnetohydrodynamic Modeling.* VACET is providing novel visual data analysis tools for use on accelerating scientific understanding of data produced by MHD codes.
- D. Batchelor, ORNL. *Simulation of Wave Interactions with Magnetohydrodynamics.* This active project has made progress in topological analysis of magnetic field lines and a data conversion utility for loading simulation data into VisIt.
- J. Cary, Tech-X. *Framework Application for Core-Edge Transport Simulations.* This potential stakeholder project is still early in its design stage and is not quite ready as of the time of this writing for a collaboration with VACET. A representative from VACET attended the FACETS project kickoff meeting in Boulder, CO in December 2006.
- P. Bonoli, MIT. *Numerical Computation of Wave-Plasma Interactions in Multi-dimensional Systems.* This prospective VACET project, which is the subject of May 2007 SAP proposal, would focus on providing fundamental visual data analysis infrastructure and expert assistance applying it to fusion data analysis, visual data analysis of phenomena in the confinement chamber: particles, field lines and scalars.
- S. Jardin, PPPL. *Center for Extended Magnetohydrodynamic Modeling.* Presently, VACET team member A. Sanderson (Utah) is on an existing SAP with CEMM. New work, which is the subject of a May 2007 SAP proposal, we would continue work on topological analysis of Poincaré plots to automatically find islands, and provide fundamental visual data analysis

infrastructure to replace aging (and no longer supported) tools presently in use at PPPL for day-to-day analysis.

- W. Nevins, LLNL. *The Plasma Microturbulence Project.* This prospective project, which is the subject of a May 2007 SAP proposal, would focus on tools and techniques for visual data analysis of particle-based data at multiple scales and field line integration.

8.2 Centers

- Dean Williams, LLNL. *Scaling the Earth System Grid to Petascale Data Center for Enabling Technologies.* Dean's team produces software infrastructure that is widely used in the climate research community for data storage, sharing, analysis and visualization. VACET's new technical contributions will be deployed through CDAT to achieve broadest possible market penetration.
- P. Colella, LBNL. *The Applied Partial Differential Equations Center for Enabling Technologies.* Colella and his team are one of VACET's primary stakeholders. VACET is working to provide a production-quality replacement for the older ChomboVis application. The impact to APDEC is a significant cost savings and the ability to perform visual analysis of large and complex AMR data on parallel machines.
- A. Shoshani, LBNL. *The Scientific Data Management Center.* We have identified several stakeholder projects where SDM's FastBit index/query technology would be helpful for accelerating visual analytics applications, and where VACET's visual analytics capabilities would be included in scientific workflows. These are "future work" items.
- D. Skinner, LBNL. *The SciDAC Outreach Center.* VACET has provided input to Skinner concerning VACET needs of the Outreach Center. We have highlighted the need for support for collaborative software and release engineering capability, as well as collaborative documents, spreadsheets and project management tools. As a result, Skinner is including GForge[2] as part of the Outreach Center's infrastructure offerings.

8.3 Institutes

- J. Owens and K.-L. Ma, UCD, for the *SciDAC Institute for Ultrascale Visualization.* Active, Joint project studying use of collections of GPUs to accelerate large data rendering.
- W. Kramer, LBNL, for the Petascale Data Storage Institute. Prospective VACET interactions would focus on joining forces with PDSI on data I/O infrastructure, models and formats for petascale applications.

8.4 Partnerships

- J. Bell, LBNL. Bell is PI of the Science Application Partnership entitled *Computational Astrophysics Consortium: Adaptive Algorithms*, which is part of Woosley's Astrophysics Consortium. VACET interactions with Bell span two different focus areas. First, VACET is working to provide to production-quality AMR visual data analysis software infrastructure to Bell's team (Section 12.6.1). Second, VACET is engaged in focused research projects that address specific scientific data understanding challenges.
- K. Stockinger, LBNL. Stockinger is Co-PI of a Science Application Partnership entitled *H5part: High Performance Data I/O for the Accelerator Community*, which is part of Ryne's concluding SciDAC Accelerator project (Section 8.1.1). By interacting with this project,

[2] http://www.gforge.org

which aims to develop and deploy an Accelerator community-wide data model and I/O capability that includes integration of index/search technology from the SDM Center, we are well positioned to quickly achieve dramatic results in the Accelerator community when the new Accelerator SciDAC projects begin in mid-2007. A follow-on Partnership entitled *High Performance Parallel Data Management and Analysis Tools for Beam Dynamics Modeling* is part of the P. Spentzouris Accelerator proposal currently under consideration for a SciDAC award. Note: this SAP project, which was included with the COMPASS Accelerator project, was not renewed in Fall 2007 due to lack of SAP funds, and as a result, this project will be soon coming to a close.

8.5 Facilities

- NERSC and its Analytics Team[3], which provides front-line visual data analysis and analytics support to projects hosted at NERSC. The scope of that Team's effort includes data management, workflow management, visualization, analysis and analytics. VACET works closely with members of NERSC Analytics in several different capacities: we share common stakeholders; NERSC Analytics is a consumer of VACET technologies and assistances; through ERCAP, NERSC provides access to computational and storage resources for use by VACET in fulfilling its production visual data understanding mission.
- The Leadership Computing Facility at ORNL provides large, open computational platforms and staff support to the DOE science community. The VACET team augments the existing Visualization team at ORNL, who provides production visualization support for data analysis and scientific understanding. Through our partnerships with ORNL users, and VACET's allocation on the ORNL INCITE compute resources, we are well poised to deploy new petascale visual data analysis capabilities to SciDAC users at ORNL.

8.6 Other Research Efforts

- LBNL Visualization Group[4], which has ongoing R&D projects in the areas of high performance Query-Driven Visualization and Analytics.
- Nearby Supernova Factory[5]. The SNfactory is an experiment to develop Type Ia supernovae as tools to measure the expansion history of the universe and explore the nature of Dark Energy. It is an international collaboration between several groups in the US and France. It is a "proving ground" for the younger Supernova Acceleration Probe (SNAP) experiment[6]. The SNfactory team will directly benefit from our work on Spectrum Synthesis analysis. One of our stakeholder projects in this subject area, P. Nugent (LBNL) for S. Woosley, is part of the SNfactory team. In addition, a VACET team member, C. Aragon (LBNL) is the team leader for the SNfactory data management and analysis infrastructure effort.

8.7 INCITE

- Warren Mori, UCLA. *Petascale Particle-in-Cell Simulations of Plasma Based Accelerators.* Prospective VACET stakeholder, joint with the NERSC Analytics team.
- Chuang Ren, Rochester. *Three-Dimensional Particle-in-Cell Simulations for Fast Ignition*, Prospective VACET stakeholder, joint with the NERSC Analytics team.

[3]http://www.nersc.gov/nusers/analytics
[4]http://vis.lbl.gov/
[5]http://snfactory.lbl.gov
[6]http://snap.lbl.gov

- Jacqueline Chen and Evatt Hawkes, SNL-CA. *Direct Numerical Simulation of Turbulent Non-Premixed Combustion – Fundamental Insight Towards Predictive Modeling.* VACET's work with Chen is described above in Section 8.1.4.

9 Change Control

Over the course of any project, changes are inevitable. Our project's mission statement – "to foster scientific insight through creating and deploying effective data understanding technology that is truly responsive to the needs of our stakeholders in the scientific research community" – will likely require constant tuning, adaptation and evolution of VACET's approach, scope, operational model. Change is inevitable in light of the fact VACET's customers will change, the technology environment will change, new opportunities will arise, DOE programmatic priorities will change, and so forth. Therefore, our overall objective here is to optimize the benefits created by VACET while minimizing the adverse impact of change. To a large extent, success in this regard will require careful management of project scope and focusing resources in areas with the greatest impact.

In this section, we discuss briefly our approach to Change Control for several different scenarios in which project change can be reasonably anticipated.

9.1 Engaging with Customers

The decision to engage or not with a particular new stakeholder project is a "business decision" that requires the approval of the project PIs. The mechanism we use for evaluating such a decision is as follows: (1) the existence of an opportunity for work with a new stakeholder is brought to the attention of the Executive Committee; (2) the potential cost/benefit of working with that stakeholder is discussed amongst the EC via email and/or at one of the regularly scheduled EC calls; (3) the EC makes a recommendation; (4) the PIs render a decision taking into account the EC's recommendation.

In some instances, there are factors external to VACET that would influence the decision to engage or not with a particular stakeholder. Such factors would include input from outside sources, like ASCR/DOE.

9.2 Scope of Customer Project

A "typical" customer/stakeholder project involves first ascertaining stakeholder needs, translating those needs into tasks or deliverables, selecting some subset of those deliverables for VACET action, performing the work to produce the deliverable, and iterate. Other aspects of this life cycle, such as long-term maintenance and support of the deliverable, are omitted from the discussion here for brevity.

Once a decision is made to proceed with a specific customer, the Customer Project Manager is responsible for generating the list of stakeholder needs and translating those needs into tasks or deliverables. The decision of whether or not to proceed with any/all of those tasks is also a "business decision" requiring EC review and recommendation and PI approval. This decision is important because it commits VACET resources, and often requires coordinating efforts across a number of different teams. Changes to scope – e.g. adding or removing deliverables – must be subject to EC review and recommendation and PI approval.

9.3 Project Management Plan

This Project Management Plan serves several purposes. It reflects the VACET mission, approach, benefits, etc. – it is a "business plan." As it is regularly updated, it is a record of accomplishments. It also presents a snapshot of goals and objectives for the future. As such, changes to "the business plan" as well as the list of future work targets requires EC review and recommendation and PI approval. The list of accomplishments are provided by all team members and reviewed by the EC. The PIs approval on the PMP is implicit due to the fact the PIs are responsible for generating the PMP with EC oversight and input from the entire team.

10 Center-Wide Accomplishments

10.1 Stakeholder Projects

In its first twelve months of operation, VACET has begun meaningful, high-impact work with stakeholder projects in a number of different science areas.

10.1.1 Accelerator Modeling

One of our stakeholder projects aims at applying high-quality interactive rendering techniques applied to particle-based datasets, such as those from PIC-based codes to perform Plasma-Wakefield accelerator modeling (Section 12.1.1). The primary stakeholder for this project is the VORPAL development team at Tech-X Corporation[7].

We have expended a great deal of effort on several related technology areas that have benefit for large, complex particle-based datasets (see Sections 12.5.2, 12.9.3, 12.5.1, 12.7.5, and 12.9.2). As our relationship with the new SciDAC Community Petascale Project for Accelerator Science and Simulation (COMPASS) matures, these technologies can be rapidly brought to bear on visual data understanding problems from the accelerator community.

We have participated in the first COMPASS All Hands Meeting in September 2007. We delivered a presentation entitled "Visualization, VACET and the SciDAC COMPASS Accelerator Project."

10.1.2 Astrophysics

As documented on the VACET wiki[8], our primary Astrophysics stakeholder is Stan Woosley's Computational Astrophysics Consortium SciDAC Science Application[9]. Our interactions with that team focus on two primary projects.

The first is providing production quality AMR visual data analysis infrastructure in support of John Bell, who is responsible for creating the AMR-based computational infrastructure for modeling Supernova explosions. VACET is achieving some economy of scale in this project by grouping together John Bell's needs with those of Phil Collela's APDEC Center, which also focuses on R&D of AMR-based computational infrastructure. The needs of these respective projects overlap to a large degree, so our overall objective is to provide production-quality AMR visual analysis infrastructure to both teams. The specific needs of these groups are documented on the VACET wiki[10].

[7] http://www.txcorp.com/products/VORPAL/
[8] http://sci.utah.edu/vacetwiki
[9] http://www.scidac.gov/physics/grb.html
[10] http://www.sci.utah.edu/vacetwiki/index.php/Collab:AMR_Applications

The second is in developing tools and techniques for replacing "Chi by eye" supernova spectrum/light curve comparative analysis. This effort has special challenges and needs, which are documented on the VACET wiki[11]. Work in this area will be applicable to future stakeholder projects where some form of machine learning/classification methods will be useful to discover relationships in large, multidimensional datasets. This particular project is groundbreaking in several respects. First, replacing "Chi by eye" with rigorous, quantitative methods will represent a major leap forward in data understanding capabilities. Second, the intersection of visual analytics with analysis and visualization on such problems represents a major growth area in our field. Third, this type of approach is promising for providing traction on Petascale and beyond data understanding in a way not possible when using only visualization or only data analysis techniques. This project's accomplishments, plans and risks are listed in Section 12.2.1.

To this end, we developed a new visual interface for rapidly exploring large collections of supernovae spectra data. This new interface, which uses well established principles of user interface design, fulfills a scientific need not met by any existing capability: the need to compare large collections of spectra over different dimensions. In the past, when only a handful of supernovae were discovered each year, supernova scientists were able to recognize spectral signatures by supernova name. However, as the SNfactory has generated a large and growing supernova spectral database, exploring the dataset for analysis has become more difficult. See Section 12.2.1 for more details.

10.1.3 Climate

VACET's multi-faceted approach to addressing the needs of the climate community is as follows. First, we are working closely with Dean Williams, PI of the Earth Systems Grid SciDAC project[12]; Dean is the lead on the Climate Data Analysis Toolkit (CDAT)[13], which is widely used in the climate community for data analysis and visualization. VACET's approach here is to achieve maximum possible market penetration into the climate research community by providing new capabilities in one of its most widely used software toolkits for data management, analysis and visualization. The second approach is to form close collaborations with specific stakeholders. At the present time, our climate stakeholder is John Drake (ORNL), who is PI on the new SciDAC project "A Scalable and Extensible Earth System Model for Climate Change Science"[14]. Both Williams and Drake have provided to VACET a set of "stakeholder needs" that form the basis for current work activities. These needs are documented on the VACET wiki[15].

Two active projects focus on (1) improving CDAT user productivity while enabling new forms of comparative visualization and analysis (Section 12.3.1); (2) deploying new advanced visualization capabilities in CDAT for use by the climate community (Section 12.3.1).

During the past year, VACET has made a breakthrough in combining advanced visualization capabilities with the CDAT infrastructure. We have developed an initial python wrapper around the ViSUS code base that is then callable from CDAT. This result enables direct access to the ViSUS scene graph and its full manipulation via python scripting. In the process we have made available fundamental features such as isosurface computation, volume rendering and computation of arbitrary slices while developed new nodes that add high quality labels to the scene (using an external font library) and tick marks for axes and color bars. The image shown in Figure 7 was generated by the python script `Cloudiness.py`. In the current interaction with the ESG group, we

[11]http://www.sci.utah.edu/vacetwiki/index.php/Collab:Astrophysics
[12]http://www.scidac.gov/compsci/ESG.html
[13]http://www-pcmdi.llnl.gov/software-portal/cdat
[14]http://www.scidac.gov/climate/earth.html
[15]http://www.sci.utah.edu/vacetwiki/index.php/Collab:Climate_Modeling

want to refine the API exposed in python since the current version is simply exposing all parameters available in the C++ implementation.

This initial success led to subsequent successes: we are now exploring alternative scene layouts for data exploration and movie creation. For example, the image in Figure 8 shows a possible layout including both embedded 3D data and slices on the side. This type of layout can increase both the realistic presentation of the information and the effective communication scientific the results. After completing our discussion with John Drake, Jamison Daniel, and the rest of the climate team at ORNL we will improve the layout and content of the scene and generate a complete animation via scripting embeddable in CDAT. This will allow for example to give the script to the users and produce similar presentation by simply changing the data source in the python script. Significance: have some initial infrastructure in place, can move forward with better vis work.

10.1.4 Combustion

VACET has launched a number of interrelated projects for combustion stakeholders. One major thrust is R&D and deployment of production quality AMR visual data analysis software infrastructure (Section 12.6.1). Others are aimed at R&D of new capabilities needed by combustion researchers to understand their data. These projects include topologically based feature identification, tracking and analysis (Sections 12.4.1 and 12.4.2); and multivariate, spatiotemporal visual data analysis 12.4.3); advanced techniques for vector field visualization (Sections 12.10.1, 12.10.2).

We implemented a new prototype multiresolution, temporal capable, topological analysis, feature detection, tracking capability for data produced by the S3D code, a candidate petascale application. This work is described in more detail in Section 12.4.2 and has resulted in numerous publications.

10.1.5 Fusion

VACET has a number of focused R&D projects, each of which is tackling a different aspect of a larger set of data understanding needs that arise in Fusion. These new techniques are being designed, implemented and tested in one or both of our production delivery platforms, SCIRun and VisIt.

One project focuses on visual analysis of particle orbits in very large particle-based datasets (Section 12.5.2). Another performs topological analysis of Poincaré plots to identify formation of islands in magnetic fields (Section 12.5.3) in MHD simulation output.

An unexpected byproduct of this topological analysis work is a new technique for multivariate particle-based visualization. We combined this work with a new query-driven visualization technique to aid in rapid visual exploration of large complex data from GTC, a candidate petascale application. This work is described in more detail in Section 12.5.2.

Yet another project explores use of linked views and a parallel coordinates display and selection mechanism to provide the ability to rapidly select subsets of large and complex data for analysis on the basis of combinations of data ranges across a number of variables (Section 12.9.2). While initially motivated by the needs of the multiple Fusion Microturbulence projects (i.e., particle-based datasets), the concepts are generally applicable to any type of large and complex dataset. It is especially noteworthy that the results of this latter R&D project were implemented in VisIt in late 2006/early 2007, integrated into the VisIt code base and made part of a public VisIt release in April 2007. During integration, testing and release, we stress-tested the new capabilities with large dataset consisting of approximately 60 million zones from the S3D code.

A challenge faced by many particle-based codes (GTC, Vorpal, etc.) is the fact that a complete data dump may not be possible nor practical. Current approaches dump every n'th particle, which may or may not be a representative sample. One of our projects focuses on developing suitable statistical sampling techniques by leveraging the R statisticspackage to help provide a better mechanism for downsampling during I/O, one that is based on a statistical model rather than one that is completely arbitrary (Section 12.9.3).

During the past several months, we have put the necessary infrastructure created and put in place at ORNL for statistical sampling research. The thesis is that large particle-in-cell based codes, like GTC, can reduce their I/O load by saving a reduced-size set of particles that has characteristics similar to the larger population, and that statistical techniques can help provide a smaller, representative population sample.

We have engaged with the FACETS project at a team-wide scale, and have developed an action plan for providing production quality visual data analysis infrastructure to the entire DOE SciDAC Fusion community (see Section 12.5.1). This effort is at risk due to the absence of SAP funding to accompany the recent new crop of Fusion SciDAC projects. While this plan is ambitious and far-reaching in effect, we have begun to take aim at "low-hanging fruit" that while may not result in research publications, will have a significant positive impact on our customer base. One such item was to resolve a client/server network-based communication issue in VisIt that often resulted in difficult use. This result is described in Section 12.7.4. Another is easing the ability to import/export data to VisIt (see Section 12.7.5).

We also authored approximately four different SAP project proposals that were submitted with Fusion Science Application proposals in the recent Fusion SciDAC CFP. While three of the four application proposals were selected for an award, there does not appear to be any available funding for ASCR SAP projects in fusion. This situation poses our project significant difficulties.

10.1.6 Mathematics

VACET's primary stakeholder in this space is the P. Colella's APDEC center. Our team is aiming to provide APDEC with a wholesale replacement for the older ChomboVis application. Our replacement, VisIt, is production-quality visual data analysis infrastructure that is suitable for use on parallel machines and on AMR data. Over the course of the next several years, we have a staged sequence of R&D plans that will result in software releases containing a lengthy set of features/capabilities needed by APDEC, as well as those APDEC support. By fostering this close relationship with APDEC, VACET will be able to achieve impact across a broader range of science applications than it would by itself.

VACET projects include: (1) a large software research, development and engineering effort that involves many of the VACET family (Section 12.6.1); (2) a subproject focusing on R&D of techniques for extraction of Embedded Boundaries in multifluid or multiphase problems on block-structured adaptive grids (Section 12.6.3); (3) a concerted effort to enable launching a visual data analysis tool from the debugger while running an AMR application (Section 12.6.2); (4) a team-wide software engineering effort to provide a library form of a state-of-the-art volume rendering engine, and integrate it into our delivery platforms (Section 12.7.1)

VACET has achieved a major breakthrough in the area of production-quality AMR visual data analysis infrastructure. Over the past year, we have undertaken a major effort aimed at helping APDEC to supplant its in-house AMR visualization tool with one that is production quality, runs on parallel machines, runs in a remote/distributed fashion, and that provides for many of its needs. In this collaboration, we first performed a gap analysis to determine a priority order of features missing from VisIt, one of our production software platforms. Most of these critical features were added to

VisIt, and as of Fall 2007, the points of contact for the Chombo team in APDEC (T. Ligocki and B. van Straalen) report that they **have transitioned to using VisIt** for the majority of the time and are poised to roll VisIt out to the entire team. In addition, many external collaborators have already begun using VisIt instead of ChomboVis. For more information, see Section 12.6.1.

10.2 Turbulence

The Livermore VACET group, developed for the first time feature based analysis of extremely high-resolution simulations of turbulent mixing created by Andy Cook and William Cabot, members of the SciDAC physics project for "Simulations of Turbulent Flows with Strong Shocks and Density Variations." The work focuses on Rayleigh-Taylor instabilities, which are created when a heavy fluid is placed above a light fluid and tiny vertical perturbations in the interface create a characteristic structure of rising bubbles and falling spikes. Rayleigh-Taylor instabilities have received much attention over the past half-century because of their importance in understanding many natural and man-made phenomena, ranging from the rate of formation of heavy elements in supernovae to the design of capsules for Inertial Confinement Fusion. This work is described in Section 12.8.1, and has resulted in numerous publications, one of which won the Best Application Paper award at IEEE Visualization.

10.3 Forward-Looking Projects

After interactions with many stakeholders, we quickly realized that there are many needs we can anticipate in the future based upon current trends in computer and computational science. To be well prepared to meet what we consider to be imminent future needs, we have undertaken a number of focused R&D projects aimed at pathfinding in a number of areas we feel will be crucial in the near term.

One project aims to explore techniques for discovery and visual display of interrelationships between variables in a large and complex dataset (Section 12.9.4).

Another is to develop scalable statistical techniques suitable for use on parallel platforms, for use in visualization as well as analysis Comparative Visualization/Analytics/Analysis (Section 12.9.5).

A trend of increasing importance is the notion of comparative visualization and analysis. Our effort here aims to move well beyond "Chi-by-eye" and A vs. B comparisons by moving into ensemble-level comparison, as well as comparison of data from different types of grids (Section 12.9.6).

One promising approach to visual analysis of highly multivariate datasets is the so-called "Function Field", where an $n-$dimensional space is mapped to a one-dimensional field (Section 12.9.7).

VACET researchers, in collaboration with researchers from the Institute for Ultrascale Visualization are working to develop a new system which enables the rendering of very complex, out-of-core scenes with unbiased global illumination. The system facilitates the rendering of hundreds of millions of triangles, gigabytes of texture, and complex shading by forcing high levels of coherency. (Section 12.10.4).

Related, our team is pursuing algorithmic R&D and software engineering for highly scalable platforms. One project focuses on techniques for implicit surface reconstruction in a way amenable to parallelization on CPU or GPU platforms. Another aims to provide the ability to perform multivariate volume visualization on GPU-based platforms using GLSL (Section 12.10.3).

10.4 Topological Analysis

As part of its research portfolio, VACET investigated new topological analysis technique for quantitative analysis of complex structures (channels) in porous media. Our new Morse theoretical approach analyzes the topology of the scalar field generated by the atomistic simulation are recover the most stable filament structures by filtration of unstable topological structures. Gives basis for analyzing structural changes that occur over time. Initially applied to nanoscale materials science simulations, our new technique holds promise for application to many other domains, including combustion. This work is described in more detail in Section 12.9.1.

10.5 Center-Wide Projects

10.5.1 Particle Data Focus Group

As described above, we have several active stakeholder projects that address the general theme of visual data analysis of particle-based datasets: Accelerator Modeling (Section 12.1.1), Fusion (Sections 12.5.2, 12.9.2). We expect to begin operations with at least two new stakeholders in the months ahead (Mori, UCLA and Ren, Rochester; both are INCITE awardees) who use a PIC code (Vorpal) to model Plasma-Wakefield accelerators; this PIC code generates large particle-based datasets.

While each of these customer/stakeholder projects has a corresponding VACET Customer Project Manager, all efforts are "coordinated" through a central VACET Particle Data Focus Group. Here, all persons working on visual analysis of particle-based datasets exchange information and ideas through conference calls and email. We look for opportunities to share ideas and technologies across these different projects.

A good example of such a shared technology, which has been implemented and deployed in one of our production delivery platforms, is the work with parallel coordinates to perform subset selection (Section 12.9.2).

10.6 Software Engineering Team

As described earlier in Section 6, VACET has a Software Engineering team comprised of representatives from all participating institutions. This team is tasked with centralized coordination of creating well-engineered software in the form of applications, libraries and modules. The team has held regular meetings and calls, and has been active in producing new software tools needed by our stakeholders.

The team has made excellent progress on a large, Center-wide activity aimed at providing a high-quality volume rendering capability for use in multiple deployment vehicles (e.g., Section 12.7.1). This new capability will have an impact in three different delivery platforms for at least six different stakeholder projects spanning diverse scientific disciplines. Initial results from this integration are shown in Section 12.7.1.

The team successfully executed a transition of the VisIt source code repository (including revision control) from one in-house at LLNL to transition to subversion server provided by the SciDAC Outreach Center. The VisIt Subversion repository is available to all VisIt developers at `svn+ssh://svn.nersc.gov/svn/visit`. This repository was made available by the SciDAC Outreach Center, who have also provided Subversion consulting support. The repository has proven very stable. Further, a VisIt release was made using Subversion, stressing branch management, etc, and everything went very smoothly. This major success reflects the quality and commitment of

both the VACET Software Engineering Team and the SciDAC Outreach Center staff who supported the effort.

10.7 Other Collaborations

In addition to the productive relationships with Science Application stakeholders, VACET has active and gestating collaborations with other SciDAC Centers, Partnerships and Institutes.

- **SDM Center.** VACET coordinates activities with the SDM Center to look for opportunities to share technology. VACET is presently evaluating several projects where the SDM Center's FastBit index/query technology could be used to accelerate visual analytics operations on large data. VACET is also coordinating with the SDM Center to join forces on at least two Fusion Science Application Partnership project proposals that will be submitted in May 2007. Areas of overlap in those projects include index/query and visual analytics as part of scientific workflows.
- **Visualization Institute.** We have had multiple positive interactions with IUSV. IUSV hosted a workshop at SC06 and invited several speakers from VACET. We also have engaged in a joint project involving use of many GPUs to address large-scale rendering problems (Section 12.10.4).
- **Petascale Data Storage Institute.** One of the focus areas of the PDSI is to improve high performance I/O for Petascale applications. Through PDSI participants at NERSC, VACET has engaged in preliminary discussions centering around the theme of working with PDSI on data I/O improvements for petascale applications. VACET's interest in this space is on data models and formats, since our infrastructure will be used for visual data analysis of data produced by the petascale applications. By becoming involved early in formulating and executing an I/O solution for those applications, VACET will be in a good position to rapidly achieve results as those applications come online and begin producing data.

10.8 Infrastructure

During the first year of operations, we have begun the process of securing access to computational and storage infrastructure for use by VACET team members in fulfilling their mission. Our primary focus thus far has been on securing access to NERSC/LBNL and LCF/ORNL for the purposes of: (1) ensuring that our technologies are operational on these platforms, since they are one of our primary deployment targets; (2) beginning to deploy our technologies at these facilities for use on stakeholder data, which tends to be primarily located at one of these two facilities.

Through the ERCAP process, we have been awarded 20,000 hours and 50 TB of storage at NERSC for AY2007. Approximately a dozen VACET members now have accounts at NERSC (all VACET members are eligible for an account at NERSC).

Through ORNL's Director's Discretionary Allocation process, we have been awarded 100,000 hours on ORNL's Jaguar XT4 system, currently the 2nd fastest HPC system in the world. This allocation inludes access to the associated storage and visualization systems. All VACET members are eligible to access this allocation at ORNL.

Joy at UCD has purchased a small amount of equipment needed to support VACET R&D activities.

10.9 Communication

A summary of our "communication" accomplishments is as follows:

- The VACET EC has engaged in regular phone conferences and email discussions to ensure coordinated management and execution of all VACET activities.
- The VACET Software Engineering team has engaged in regular calls and email exchanges to coordinate and execute Center-wide software projects.
- The VACET Particle focus group has been active in coordinating activities related to visual data analysis of particle-based datasets.
- VACET team members have participated in dozens of face-to-face meetings with active and prospective stakeholders and collaborators.
- The VACET website, www.vacet.org, was brought online and has been undergoing regular updates.
- The VACET wiki[16] has been brought online and is an indispensable part of our Center's operations.
- Two All-Hands meetings. The first was in October 2006 in conjunction with IEEE Visualization 2006. The second was in late January 2007 at Salt Lake City, UT.
- HPCwire interview[17] with VACET personnel.

10.10 Publications

Refereed Journal Papers

1. *December 2007*, S. Dillard, V. Natarajan, G. Weber, V. Pascucci, and B. Hamann. "Topology-guided tessellation of quadratic elements," *International Journal of Computational Geometry and Applications (IJCGA)*, to appear. [20].
2. *December 2007* A. Sanderson, M. Meyer, R. Kirby, and C. Johnson. A framework for exploring numerical solutions of advection reaction diffusion equations using a gpu based approach. *Journal of Computing and Visualization in Science*, 2007. to appear [48].
3. *December 2007* J. Shepherd and C. Johnson. Hexahedral mesh generation constraints. *Engineering with Computers*, 2007. to appear [51].
4. *October 2007*, L.J. Gosink, J.C. Anderson, and K.I. Joy, "Variable Interactions in Query Driven Visualization," IEEE Transactions on Visualization and Computer Graphics (Proceedings of IEEE Visualization 2007), 13(6), October 2007 [35].
5. *October 2007*, A. Gyulassy, M. Duchaineau, V. Natarajan, V. Pascucci, E. Bringa, A. Higginbotham, and B. Hamann; "Topologically Clean Distance Fields," IEEE Transactions on Visualization and Computer Graphics 13(5), Proceedings of IEEE Visualization 2007 [24].
6. *October 2007*, A. Gyulassy, V. Natarajan, B. Hamann, and V. Pascucci; "Efficient Computation of Morse-Smale Complexes for Three-dimensional Scalar Functions," IEEE Transactions on Visualization and Computer Graphics 13(5), Proceedings of IEEE Visualization 2007 [6].
7. *October 2007*, C. Garth, F. Gerhardt, X. Tricoche, H. Hagen. "Efficient Computation and Visualization of Coherent Structures in Fluid Flow Applications," IEEE Transactions on Visualization and Computer Graphics 13(5), Proceedings of IEEE Visualization 2007 [22].
8. *October 2007*, Carlos Scheidegger, David Koop, Huy Vo, Juliana Freire, and Claudio Silva. "Querying and Creating Visualizations by Analogy." IEEE Transactions on Visualization and Computer Graphics 13(5), Proceedings of IEEE Visualization 2007 [49]

[16] http://www.sci.utah.edu/vacetwiki/index.php/Main_Page
[17] http://www.lbl.gov/CS/Archive/news110306.html

9. *October 2007*, I. Wald, H. Friedrich, A. Knoll, C.D. Hansen. "Interactive Isosurface Ray Tracing of Time-Varying Tetrahedral Volumes." IEEE Transactions on Visualization and Computer Graphics 13(5), Proceedings of IEEE Visualization 2007 [54]
10. *July 2007*, V. Pascucci, G. Scorzelli, P.-T. Bremer, and A. Mascarenhas, "Robust On-line Computation of Reeb Graphs: Simplicity and Speed" ACM Transactions on Graphics (SIGGRAPH 2007), 26(3), July 2007 [47].
11. *October 2007*, G. H. Weber, P.-T. Bremer, V. Pascucci. "Topological Landscapes: A Terrain Metaphor for Scientific Data," IEEE Transactions on Visualization and Computer Graphics 13(5), Proceedings of IEEE Visualization 2007 [57] .
12. *October 2007*, S. Yoon and P. Lindstrom. "Random-Accessible Compressed Triangle Meshes." IEEE Transactions on Visualization and Computer Graphics 13(5), Proceedings of IEEE Visualization 2007 [59].
13. *October 2007*, A. Wiebel, X. Tricoche, D. Schneider, H. Jänicke, and G. Scheuermann. "Generalized Streak Lines: Analysis and Visualization of Boundary Induced Vortices." IEEE Transactions on Visualization and Computer Graphics 13(5), Proceedings of IEEE Visualization 2007 [58].
14. *January 2008*, J. Chen, I. Yoon and E. W. Bethel. "Interactive, Internet Delivery of Visualization via Structured, Prerendered Multiresolution Imagery." IEEE Transactions on Visualization and Computer Graphics, Jan/Feb 2008 (to appear) [15].
15. *June 2007*, S. Ahern, "Petascale Visual Data Analysis in a Production Computing Environment," Journal of Physics, Conference Series – SciDAC 2007, 78, 2007, Boston, MA, USA. [3]
16. *June 2007*, E.W. Bethel, C.R. Johnson, K. Joy, S. Ahern, V. Pascucci, H. Childs, J. Cohen, M. Duchaineau, B. Hamann, C. Hansen, D. Laney, P. Lindstrom, J. Meredith, G. Ostrouchov, S.G. Parker, C.T. Silva, A. Sanderson, X. Tricoche. "SciDAC Visualization and Analytics Center for Enabling Technology," In Journal of Physics, Conference Series, 78, 2007, Boston MA, USA [11]
17. *June 2007*, P.-T. Bremer, E. M. Bringa, M. Duchaineau, A. G. Gyulassy, D. Laney, A. Mascarenhas, and V. Pascucci. "Topological Feature Extraction and Tracking," Journal of Physics, Conference Series – SciDAC 2007, 78, 2007, Boston, MA, USA [12]
18. *June 2007*, Hank Childs. "Architectural Challenges and Solutions for Petascale Postprocessing." Journal of Physics, Conference Series – SciDAC 2007, 78, 2007, Boston, MA, USA [16].
19. *June 2007*, Kenneth I. Joy, Mark Miller, Hank Childs, E. Wes Bethel, John Clyne, George Ostrouchov, and Sean Ahern. "Frameworks for Visualization at the Extreme Scale." Journal of Physics, Conference Series – SciDAC 2007, 78, 2007, Boston, MA, USA [31].
20. *June 2007*, C. Jones, K.-L. Ma, A. Sanderson, L. Myers. "Visual Interrogation of Gyrokinetic Particle Simulations," In Journal of Physics, Conference Series – SciDAC 2007, 78, 2007, Boston MA, USA [30].
21. *March/April 2007*, G. Weber, S.E. Dillard, H. Carr, V. Pascucci, and B. Hamann, "Topology-Controlled Volume Rendering", IEEE Transactions on Visualization and Computer Graphics Vol. 13, No. 2, pp. 330-341, March/April 2007 [55].
22. *January/February 2007*, H. T. Vo, S. P. Callahan, P. Lindstrom, V. Pascucci, and C. T. Silva. Streaming simplification of tetrahedral meshes. *IEEE Transactions on Visualization and Computer Graphics*, 13(1):145–155, 2007 [53].

23. *November 2007*, D. Laney, A. Mascarenhas, P. Miller, P. T. Bremer, and V. Pascucci. Understanding the Structure of the Turbulent Mixing Layer in Hydrodynamic Instabilities. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):1053–1060, 2006 [33].
24. *November 2007*, S. P. Callahan, L. Bavoil, V. Pascucci, and C. T. Silva. Progressive volume rendering of large unstructured grids. *IEEE Transactions on Visualization and Computer Graphics*, 13(1):1307–1314, 2007 [14].

Refereed Papers in Conference Proceedings

1. *November 2007*, J. Bennett, V. Pascucci, and K. Joy. Genus oblivious cross parameterization: Robust topological management of intersurface maps. In *Proceedings of Pacific Graphics 2007*, 2007. To appear [9].
2. *July 2007*, S. E. Dillard, V. Natarajan, G. H. Weber, V. Pascucci, and B. Hamann. ¡¡¡¡¡¡¡ .mine Tessellation of quadratic elements. In *ISAAC*, pages 722–731, 2006 [21].
3. *June 2007*, G. Weber, V. Beckner, H. Childs, T. Ligocki, M. Miller, B. van Straalen, E. W. Bethel, "Visualization Tools for Adaptive Mesh Refinement Data." This survey paper, authored jointly by representatives from VACET and APDEC, sketches the landscape of current AMR visualization capabilities in various tools, with a focus on the special visualization challenges posed by AMR data [56].
4. *March 2007*, C. Garth, G.-S. Li, X. Tricoche, and C. Hansen. "Interactive Visualization of Coherent Structures in Transient Flows." In Topology-based Methods in Visualization, Mathematics + Visualization (TopoInVis), March 2007, Grimma Germany [23].
5. *May 2007*, J.C. Anderson, L.J. Gosink, M.A. Duchaineau, and K.I. Joy. "Feature Identification and Extraction in Function Fields," Proceedings of EuroVis 2007, pp 195–201, May 2007, Norrköping, Sweden [4]
6. *December 2006*, P. Miller, P.-T. Bremer, W. Cabot, A. Cook, D. Laney, A. Mascarenhas, and V. Pascucci. Application of morse theory to analysis of rayleigh-taylor topology. In *10th International Workshop on the Physics of Compressible Turbulent Mixing*, 2006 [38].

Book Chapters

1. *August 2007*, S. G. Parker, K. Damevski, A. Khan, A. Swaminathan, and C. R. Johnson. *Advanced Computational Infrastructures for Parallel/Distributed Adaptive Applications*, chapter The SCIJump Framework for Parallel and Distributed Scientific Computing. Wiley Press, 2007. to appear [46].
2. *August 2007*, P.-T. Bremer and V. Pascucci. "Topology-Based Methods In Visualization", chapter A Practical Approach to Two-dimensional Scalar Topology. Springer Verlag, August 2007[13].

Edited Books

1. *November 2007*, A. Pang, C. Hansen, and M. Chen, editors. "IEEE Visualization Conference and IEEE Information Visualization Conference Proceedings 2007" IEEE Transactions on Visualization and Computer Graphics November/December 2007 (Vol. 13, No. 6) [40].
2. *November 2006*, G. Bebis, R. Boyle, B. Parvin, D. Koracin, P. Remagnino, A. V. Nefian, M. Gopi, V. Pascucci, J. Zara, J. Molineros, H. Theisel, and T. Malzbender, editors. *Advances in Visual Computing, Second International Symposium, ISVC 2006, Lake Tahoe, NV, USA, November 6-8, 2006 Proceedings, Part I*, volume 4291 of *Lecture Notes in Computer Science*. Springer, 2006 [7].

3. *November 2006*, G. Bebis, R. Boyle, B. Parvin, D. Koracin, P. Remagnino, A. V. Nefian, M. Gopi, V. Pascucci, J. Zara, J. Molineros, H. Theisel, and T. Malzbender, editors. *Advances in Visual Computing, Second International Symposium, ISVC 2006, Lake Tahoe, NV, USA, November 6-8, 2006 Proceedings, Part II*, volume 4292 of *Lecture Notes in Computer Science*. Springer, 2006 [8].

Invited Papers and Articles

1. E. Wes Bethel, Chris Johnson, Cecilia Aragon, Prabhat, Oliver Rübel, Gunther Weber, Valerio Pascucci, Hank Childs, Peer-Timo Bremer, Brad Whitlock, Sean Ahern, Jeremy Meredith, George Ostrouchov, Ken Joy, Bernd Hamann, Christoph Garth, Martin Cole, Charles Hansen, Steven Parker, Allen Sanderson, Claudio Silva, Xavier Tricoche. "DOE's SciDAC Visualization and Analytics Center for Enabling Technologies – Strategy for Petascale Visual Data Analysis Success." To appear in CTWatch Quarterly, Volume 3, Number 4, November 2007. `http://www.ctwatch.org`.

Papers Submitted or Under Review

1. H. Childs, S. Ahern, J. Meredith, M. Miller, and K.I. Joy, "Comparative Visualization Using Cross-Mesh Field Evaluations and Derived Quantities," *IEEE Transactions on Visualization and Computer Graphics*, submitted.
2. M. Miller, H. Childs, J. Clyne, G. Ostrouchov, S. Ahern, and K.I. Joy, "Equivalence Class Functions: A New and Versatile Framework for Visualization at the Extreme Scale," IEEE Visualization 2007, submitted.

   This list is not complete. VACET team members collectively made approximately 10-12 submissions to IEEE Visualization 2007. We were unable to collect data about all of these submissions in time for inclusion in this PMP.

10.11 Outreach, Presentations, Education and Service

1. *Presentation* E. Wes Bethel, *Query-Driven Visualization Accelerates Scientific Insight.* Presented at The National Science Foundation (NSF) and Cyber-enabled Discovery and Innovation (CDI) Workshop, Mathematical Sciences Research Institute (Berkeley, CA), Oct 12, 2007.
2. *Presentation* E. Wes Bethel, *Visualization, VACET and the SciDAC COMPASS Accelerator Project*, presented at Community Petascale Project for Accelerator Science and Simulation (COMPASS) All Hands Meeting, Fermi National Laboratory, Sept 17, 2007.
3. *Presentation* E. Wes Bethel, *Occam's Razor and Petascale Visual Data Analysis*, presented at 2007 Fall Creek Falls Conference, Nashville, TN, Sept 26, 2007.
4. *Presentation* Hank Childs, *Architectural Challenges and Solutions for Petascale Postprocessing*, presented at 2007 Fall Creek Falls Conference, Nashville, TN, Sept 26, 2007.
5. *Presentation* V. Pascucci, *Robust Extraction and Tracking of Topological Features in Scientific Data: State of the Art and Future Challenges*, presented at IUSV Workshop on Feature Extraction and Tracking (FET 2007), Davis, CA, August 1-2, 2007.
6. *Presentation.* V. Pascucci, *On-line computation of Reeb-graphs: simplicity and speed* Technical Paper presentation, Siggraph 2007, San Diego CA, August 2007.
7. *Presentation.* Hank Childs. *Architectural Challenges and Solutions for Petascale Postprocessing.* SciDAC 2007, Boston MA, USA.

8. *Presentation.* S. Ahern, *Petascale Visual Data Analysis in a Production Computing Environment*, SciDAC 2007, Boston, MA, USA.
9. *Presentation.* C. Johnson, *SciDAC Visualization and Analytics Center for Enabling Technology*, SciDAC 2007, Boston, MA, USA.
10. *Presentation*, Kenneth I. Joy, *Visualization and Analytics Advances for SciDAC Science,* SciDAC 2007, 2007, Boston, MA, USA.
11. *Presentation*, G. Weber, *Visualization Tools for Adaptive Mesh Refinement Data*, presented at at the 2nd International Conference on Numerical Modeling of Space Plasma Flows ASTRONUM-2007 (Paris, France, June 11-15, 2007) and the 4th High-End Visualization Workshop (University Center Obergurgl, Tyrol, Austria, June 17-22, 2007).
12. *Poster.* C. Jones, K.-L. Ma, A. Sanderson, and L. Myers. *Visual Interrogation of Gyrokinetic Particle Simulations*, 2007, Boston, MA, USA
13. *Poster.* P.-T. Bremer, *Topological Feature Extraction and Tracking*, 2007, Boston, MA, USA
14. *Service.* VACET team members serve in a leadership role for IEEE Visualization 2007. K. Joy is the General Chair. C. Hansen, E. Wes Bethel and V. Pascucci are on the Program Committee.
15. *Service.* SciDAC 2007 Program Meeting. E. Wes Bethel is on the Organizing Committee for this year's SciDAC 2007 Program meeting.
16. *Service.* ASCR Visualization and Analytics Workshop 2007. Chris Johnson and E. Wes Bethel are among the co-organizers of the ASCR Visualization and Analytics Base Program Workshop 2007.
17. *Service/Presentation.* X. Tricoche delivered a presentation entitled *Characterizing the Topology of a Hamiltonian System Exhibiting Chaos: From the Standard Map to the Tokamak* at the University of Kaiserslautern.
18. SciDAC Review Journal – Spring 2007. Several images appear in the Spring 2007 issue of the SciDAC Review Journal, including the image on the journal cover (see Figure 2). These images were produced by C. Siegerist, LBNL, in a collaborative effort between VACET and the NERSC Analytics team in support of science stakeholders with INCITE awards hosted at NERSC.
19. *Presentation.* C. Johnson, *Visual Computing: Research Challenges*, Harvard University, Boston, May, 2007.
20. *Presentation.* C. Johnson, *Large-Scale Bioimaging and Visualization*, IEEE International Parallel and Distributed Processing Symposium (IPDPS), Long Beach, March 2007 (Keynote Speaker).
21. *Presentation.* C. Johnson *Visualizing the Future*, Utah Computer Society, Salt Lake City, February 2007 (25th-Anniversary Keynote Presentation).
22. *Computational Bioimaging and Visualization: Challenges and Opportunities*, Center for Computational Molecular Biology, Brown University (Distinguished Lecture), December 2006.
23. *Visualizing the Future*, Second International Symposium on Visual Computing (ISVC06), Lake Tahoe, November 2006 (Conference Banquet Speaker).
24. *Visual Computing: Research Challenges*, Harvard University, Boston, May, 2007.
25. *Inverse Bioelectric Field Problems*, Inverse Days 2006, Tampere, Finland, December 2006.
26. *Visualizing Uncertainty*, Uncertainty Workshop, Society of Exploration Geophysicists Annual Meeting 2006, New Orleans, October 2006.
27. SciDAC 2007 Project Meeting in Boston, MA. VACET team member Bethel served on the SciDAC 2007 Organizing Committee. In the SciDAC 2007 program, from VACET there were four podium speakers, two posters, and a total of six papers.

This list is not complete.

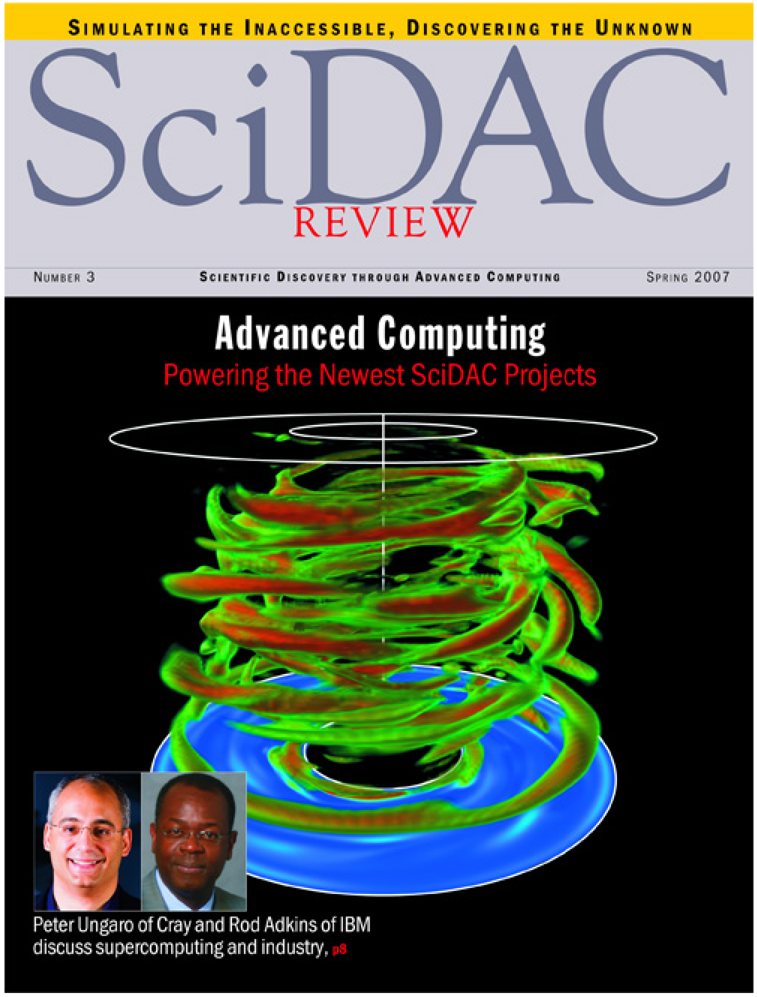

Figure 2: The cover of the Spring 2007 issue of the SciDAC Review highlights an image created in a collaborative effort between VACET and the NERSC Analytics team. It shows one time step of the evolution of total advective radial flux of axial moment, a study important in aiding understanding of the science of star and black hole formation.

10.12 Staffing Changes

One of the primary challenges with any new project is the potentially long lead time between when funding begins and teams are able to fully staff up. The following list summarizes staffing changes at each of the institutions over the past twelve months:

- **LBNL:** Prabhat[18], formerly a research associate at Brown University, is the most recent addition to the LBNL team. His broad background enables him to contribute to a number of different projects that span both VACET and NERSC Analytics.
- **LLNL:** No changes
- **ORNL:** David Pugmire joins the ORNL team.
- **Utah:** No changes
- **UC Davis:** Christoph Garth joins the UCD team as new full-time staff. He recently finished his PhD at the University of Kaiserslautern.

[18] http://vis.lbl.gov/~prabhat

11 Center-Wide Plans

11.1 Stakeholder Projects

Each Customer Project Manager has engaged in a substantial amount of project planning activity. The details of each project's planned work activities are listed below under Section 12.w

11.2 Other Collaborations

- VACET team members are contributing to five different Science Application Partnership proposals, four to Fusion and one to Accelerator. The scope of new work activities will be refined depending upon which of these are awarded and to what level.

11.3 Infrastructure

- VACET will submit a renewal ERCAP request for continued access to NERSC facilities in AY2008.
- ORNL/LCF. We anticipate being granted access to the large computational facilities at ORNL/LCF in the near future. Effort is presently underway to arrange such access.
- Per-site facilities requirements to conduct basic research and preliminary tests will be evaluated on a case-by-case basis.
- The SciDAC Outreach Center is providing access to a Subversion server so we can upload and distributed VisIt source code, track downloads, etc. (see Section 12.7.3). Additionally, we will evaluate use of GForge's collaborative project management software for use within VACET in the upcoming period.

11.4 Communication

We anticipate no substantive change in our existing methods for internal Center-wide communication. We anticipate a twice-yearly All-Hands meeting, with the next one to occur in conjunction with IEEE Visualization 2007.

One potential new avenue for communication will be the GForge server at the SciDAC Outreach Center. GForge includes collaborative project management capabilities. We will evaluate use of that capability for coordinated management of VACET projects.

We categorize the Project Management Plan as being part of the general subject of "Communication." While subject to change, our objective will be to submit an updated Project Management Plan on a twice-yearly basis. This frequency will serve to provide regular updates of progress to DOE as well as document ongoing and effective project management oversight across all of VACET's many different activities. The dates of the twice-yearly PMPs will be near the beginning and middle of each fiscal year. Therefore, the next PMPs will be submitted around 1 April 2008 and 1 October 2008.

11.5 Publications

While any prediction of publications is somewhat imprecise, we can anticipate that, roughly speaking, each stakeholder project is likely to produce one or more publications. Taking into account all stakeholder projects, we can reasonably anticipate on the order of 15-30 publication submissions over the course of the next year or so. These would span a potentially diverse range of topics: results of basic research in visualization and analytics algorithms, applications of techniques to solve domain-specific problems, management reports, web pages, movies and images.

11.6 Outreach, Presentations, Education and Service

- 2009 TopoInVis Workshop. V. Pascucci and X. Tricoche are co-charis of the 2009 edition of the TopoInVis workshop (see http://topoinvis.org).
- SC2007 Workshop. Joint with the Visualization Institute, we have organized a workshop at SC2007.
- SciDAC 2007 Program Meeting – Tutorial/Workshop. VACET will conduct a one-day workshop in conjunction with the SciDAC 2007 Program meeting in late June 2007 in Boston, MA. The objective will be to provide training in use of VACET's production visual data analysis software infrastructure to current and potential stakeholders as well as other interested parties.
- VACET team members serve in a leadership role for the 2007 prestigious Dagstuhl Seminar on "Scientific Visualization," by invitation only. K. Joy is part of the organizing committee and C. Johnson, C. Hansen, V. Pascucci, C. Silva, have been invited to give a presentation.
- International Summer School on Scientific Visualization 2007. C. Hansen, and V. Pascucci will give a cycle of lectures for a week at the summer school "Ecoles D'Ete" on the topic of "Advanced Methods in Scientific Visualization" 2007.
- IEEE Visualization 2007. VACET team members occupy many positions within the IEEE Visualization 2007 conference team. K. Joy is the conference chair. C. Hansen leads the Technical Program Committee. W. Bethel, V. Pascucci and X. Tricoche are a members of the Technical Program Committee. Most VACET members are Technical Paper reviewers.
- Chris Johnson served as Guest Co-Editor, Engineering with Computers, Special Issue on Computational Bioengineering, 2007.
- Chris Johnson served on the Scientific Advisory Board for the Institute for Computational Engineering and Science, University of Texas, Austin, April 2007.
- Chris Johnson served on the External Advisory Board for the NIH National Center for Microscopy and Imaging Research, UCSD, April 2007.
- Chris Johnson served on the External Advisory Board for the NIH National Center for Image Guided Therapy, Harvard University, June 2007.
- Chris Johnson served on the International Program Committee, International Symposium on Volume Graphics 2007.
- Chris Johnson served on the Program Committee, IEEE International Symposium on Computer-Based Medical Systems 2007.
- Chris Johnson served on the Organizing Committee, 4th International Conference on Functional Imaging and Modeling of the Heart, 2007.
- Chris Johnson served on the Scientific Advisory Board for the NIH Simbios National Center for Biomedical Computing, Stanford University, October, 2007.
- Valerio Pascucci served on the Program Committee of Eurographics 2008.
- Valerio Pascucci served on the Program Committee of the ACM Symposium on Computational Geometry (SoCG) 2007.
- Valerio Pascucci served on the Program Committee of the IEEE Conference on Shape Modeling and Applications (SMI) 2007.
- Valerio Pascucci served on the Program Committee of the International Symposium on Visual Computing (ISVC) 2007.
- Valerio Pascucci served on the Program Committee of the IASTED Conference on Graphics and Visualization in Engineering (VGE) 2007.
- Valerio Pascucci served on the Program Committee of the IASTED Conference on Visualization, Imaging, and Image Processing (VIIP) 2007.

12 Per-Project Accomplishments and Plans

The subsections that follow convey accomplishments, plans, and risks on a per-project basis.

12.1 Accelerator/HEP

12.1.1 High Quality, Interactive Rendering of Particle-Based Datasets

VACET Team: S. Parker, Utah (Team Leader) and Thiago Ize, Utah.

Customer/stakeholder: Paul Hamill and Peter Messmer at Tech-X Corporation.

Stakeholder need(s): Tech-X has previously showed the open-source POV Ray ray tracer to produce high quality ray traced visualizations of both fusion and accelerator data. We set out to demonstrate that these images could be rendered at multiple frames per second, for both interactive visualization and high quality publication images. In this context, visual cues such as shadows (including soft shadows), high quality lighting, and even reflections can help convey spatial relationships effectively.

Our initial results show that our approach is capable of generating these high-quality frames at a rate ranging from 10s to 100s of times faster than POVray, which means this technology is practical for interactive visual data analysis.

Milestones/deliverables:

- Demonstrate interactive rendering of particles similar to the offline renderings already performed at Tech-X[19] of output from the Vorpal code.
- Incorporate some of the particle binning algorithms so that the heightfield visualizations can be performed interactively as well.
- Install Manta at Tech-X.

Task Dependencies. The tasks are thus far largely independent of one another.

Accomplishments.

We started work on this project in February and we've accomplished all of the original goals. Tech-X originally wanted a way to move lights interactively so that they could see how high off the ground their particles were based off of the shadows. In the end it turned out that we could also do many more things for them than they expected. We gave them soft shadows that do an even better job (although it's not yet fast enough for interactivity since it's a temporary solution.) We showed how we can do animations (with interpolation if wanted) in real-time. We showed that we can even handle their other data sets, such as their Vorpal cavity datasets[20], which require "complex geometry". For an example of our work, see Figure 3 below as well as link on the VACET wiki[21].

We also implemented a crude user interface for recording camera paths and stopping/starting the animation when the user moves the mouse (keyframing), and added support for particle culling and faster soft shadows.

In the end Peter seemed happy with the work and said "I think it's obvious that you have all the building blocks in place for our viz needs in VORPAL."

Work Targets Next Six Months.

[19] http://fusion.txcorp.com/∼messmer/vorpal/
[20] http://fusion.txcorp.com/∼messmer/cavities/
[21] see: http://www.vacet.org/gallery/index.html#accelerator

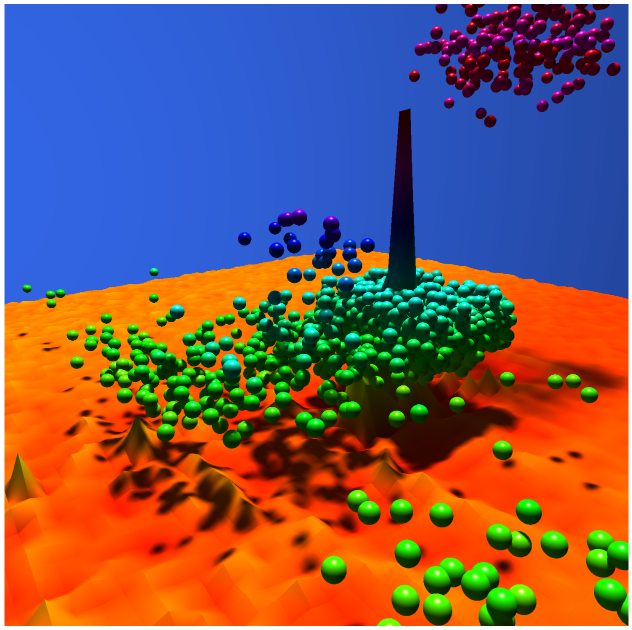

Figure 3: The heightfield represents particle density from a VORPAL dataset of a laser wakefield acceleration. The spheres are individual particles whose velocity exceeds a certain threshold. Since we only had access to a single timestep of data, we randomly perturbed the heightfield and spheres to show we can handle time-varying datasets. T. Ize, SCI. Data data provided by Tech-X corporation and Peter Messmer.

The primary task is to follow up with Tech-X to see if the have been able to use it, and if they have been able to complete their anticipated runs.

1. Implement path tracing, ambient occlusion, and material models for higher quality images.
2. Apply Manta to the data produced by their upcoming runs.
3. Install Manta at Tech-X.
4. Incorporate improved material models to further enhance the visual quality of the images.

Next 12-24 months:

If they find this tool useful, the primary improvement would be an improved user interface. The python-based user interface in Manta would allow us to quickly create a custom UI for their applications if necessary.

12.2 Astrophysics

12.2.1 Spectrum Synthesis Visual Analytics

VACET Team Members: Cecilia Aragon, LBNL (Team Leader); Ali Pinar (LBNL) and Daniela Ushizima (LBNL)

Stakeholder: Peter Nugent working with Stan Woosley on the SciDAC Science Application "Computational Astrophysics Consortium"[22].

[22] http://www.scidac.gov/physics/grb.html

High Level Stakeholder Needs: Supernova spectral analysis via visual analytics and interactive visualization. See the VACET wiki[23] for more information.

Milestones/Deliverables for Next 6 Months

Start date on this project was January 2007. A1 is ongoing. Expected completion dates.

Task 1. Research and evaluate various approaches to spectral data visualization. Expected completion: 5/31/2007.

Task 2. Select best approach(es). Estimated start and completion dates: 6/1/2007, 6/30/2007.

Task 3. Prototype visual analytics tool for spectral analysis and classification. Estimated start and completion dates: 7/1/2007, 9/15/2007.

Task 4. Apply tool to experimental supernova spectral data set. Estimated start and completion dates: 9/16/2007, 11/1/2007.

Task 5. Evaluate existing spectral data set and formats. Scientists' data is currently not flux-calibrated, but we can become familiar with data formats and files used for supernova spectra.

Task 6. Evaluate degree of noise in the data and work with scientists to model the problem. Many standard machine learning techniques do not take noise into account. We need to model it and evaluate methods of incorporating error propagation into the final feature set.

Task 7. After delivery of flux-calibrated SN spectral data set, apply statistical learning techniques to develop feature set based on scientists' needs. Scientists need a minimal feature set which will ideally fully describe key properties of the supernova, such as peak brightness. The discovery of such features will enable design of filter sets for future telescopes such as the Large Synoptic Survey Telescope (LSST)[24].

Task 8. Develop initial version of visual analytics software for SN spectral classification (will include visual user interface and machine learning algorithms). This will be based on scientists' feedback to the three prototypes developed in the previous time period.

Accomplishments

- **Task 1.** Research and evaluate various approaches to spectral data visualization. Meetings with scientists confirm that they are dealing with a larger data set than ever before and are looking for ways to help them organize and classify spectral data, and are looking for features within the spectral data that are highly correlated with important supernova information such as peak brightness. For more information, see the project's webpage[25] Completed 5/31/07.
- **Task 2.** Select best approach(es). The selection was based on meetings and feedback with project scientists. Completed 6/30/2007.
- **Task 3.** Prototype visual analytics tool for spectral analysis and classification. We built three prototypes and obtained feedback from scientists on each one. Completed 9/15/2007.
- **Task 4.** Apply tool to experimental supernova spectral data set. Tool was applied to actual supernova data set. See Figures 4, 5 and 6. Completed 10/1/2007.

Work Targets for the Next Six Months

A5 and A6 are ongoing. Expected completion dates:

[23] http://www.sci.utah.edu/vacetwiki/index.php/Collab:Astrophysics

[24] http://www.lsst.org/lsst_home.shtml

[25] http://vis.lbl.gov/Research/SpectraVis/index.html.

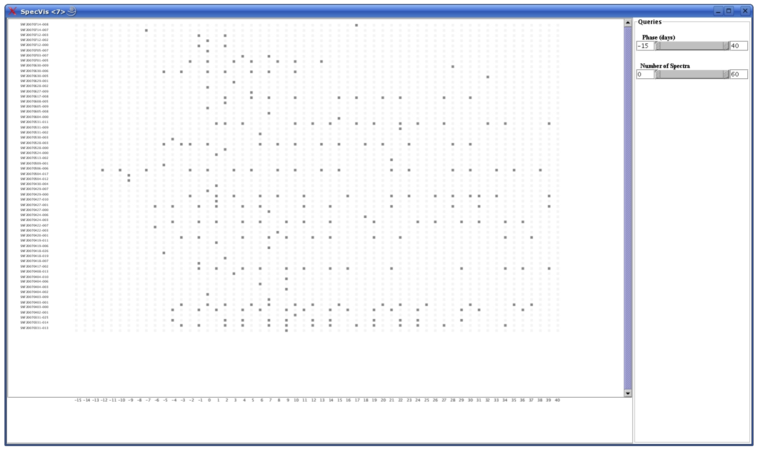

Figure 4: Our prototype implementation uses a "fisheye" style browser as an interface that enables rapid exploration of supernovae spectra obtained via observation. In the "zoomed out" view, the vertical axis contains the target names of the supernovae and on the horizontal has the age in days of the supernova or the phase. The period of scientific interest in the study of type Ia SNe ranges between 15 days before peak brightness and 40 days after peak brightness.

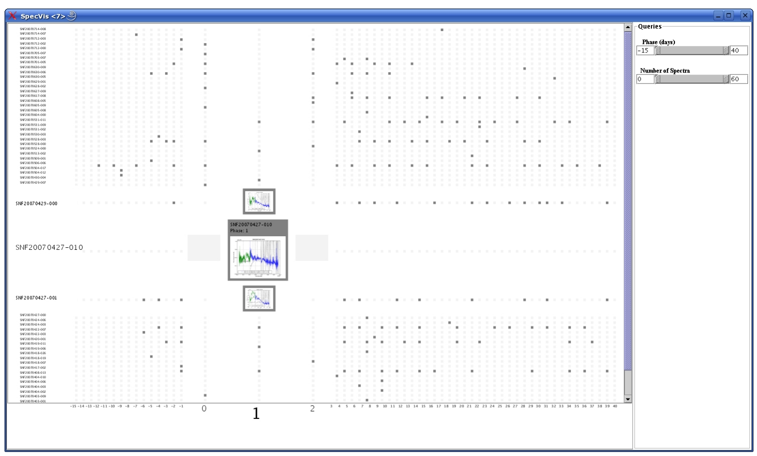

Figure 5: In order to see a more detailed spectrum, the user clicks on one of the spectrum nodes and the fisheye technique makes the clicked dot and its neighbors presented in a slightly larger way.

- A5: Nov 15, 2007
- A6: Feb 1, 2008
- A7: Apr 1, 2008 (assuming flux-calibrated data set delivered on time Nov 15, 2007)
- A8: Apr 1, 2008

A5: start date Oct 1, 2007; expected duration 1 1/2 months. Potential risks: low. Inability to meet with customer due to scheduling conflicts. Possible wrong choice of approach.

A6: start date Oct 1, 2007; expected duration 4 months. Potential risks: moderate. Modeling noisy data is a challenging problem. Inability to meet with customer due to scheduling conflicts. Possible wrong choice of approach.

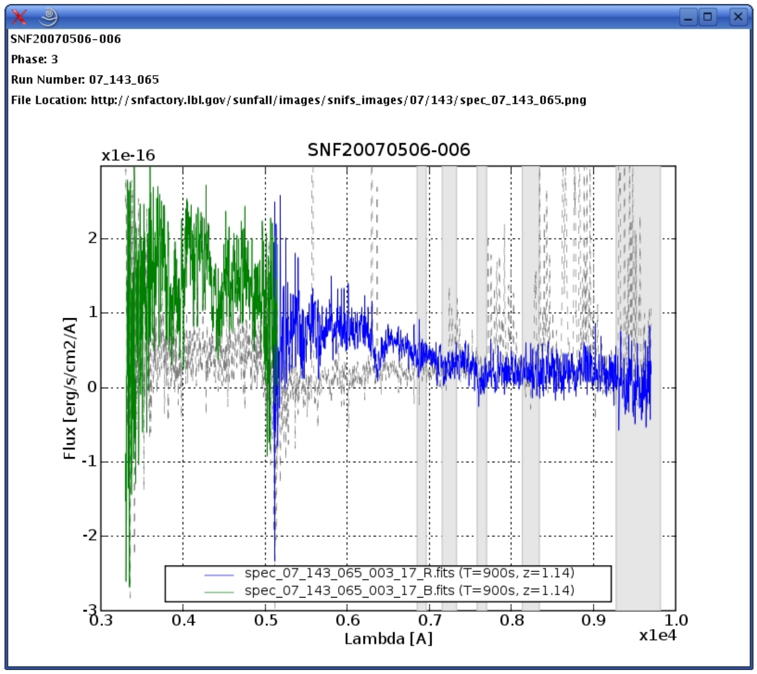

Figure 6: The user can click one more time on the spectrum of interest and the zooming is taken at a higher level. The semantic zooming is taking place at four levels: the initial dots, the two small spectra of focus and neighbor size, and the enlarged graph.

A7: start date Nov 15, 2007; expected duration 4 1/2 months. Potential risks: moderate. Major risk is delay in delivery of flux-calibrated data set from scientists. Once data delivery is complete, risks are low.

A8: start date Nov 1, 2007; expected duration five months. Potential risks: moderate. Major risk is insufficient software development staff to integrate machine learning output with visual user interface. Chosen approaches might not work well. Software performance might be insufficient. Inability to integrate with existing supernova spectral database. Assistance from experts unavailable due to scheduling conflicts.

Plans for Next 12 to 24 months:

- Iterative design applied to visual analytics tool to improve (based on scientists' feedback) feature detection and classification via machine learning and clustering algorithms applied to the spectral datasets.
- Spectra parameter fitting.
- Temporal analysis of supernova spectra.

Start date could be as early as April 2008. Estimated duration of these three tasks: 12-24 months.

The greatest project risk is lack of personnel with production programming experience.

12.3 Climate

12.3.1 Production Climate Visual Data Analysis Infrastructure

Background VACET committed to supporting the needs of the Community Climate System Model (CCSM) Consortium[26] in collaboration with the Earth System Grid[27].

Mission. The most advanced climate modeling systems seek to enable a new deeper understanding of the dynamics of global carbon cycle, atmospheric chemistry, land and ocean ecological processes and their coupling with climate. This will allow pursuing reliable answers to fundamental questions related to climate variability and global change at time scales ranging from decades to centuries. In this effort VACET will work in close collaboration with the Earth Science Grid and provide new advanced data analysis and visualization tools to the CCSM Consortium and the climate modeling community in general. One target will be the deployment of a first set of tools by FY08, in time to facilitate the analysis of data for the 5th assessment report of the Intergovernmental Panel on Climate Change (IPCC)[28].

Customers/Stakeholders:

- Community Climate System Model Consortium, Primary contacts John Drake (ORNL) and Phil Jones (LAN). Drake and Jones and co-PIs of the SciDAC Science Application entitled "Modeling the Earth System."
- Earth System Grid. Primary contact Dean Williams, PI of the SciDAC Center for Enabling Technology entitled "Sharing a World of Data – Scaling the Earth Systems Grid to Petascale Data to enable faster, easier sharing of climate change research data."
- Primary target software tools: CDAT, VCDAT. New VACET capabilities will be layered atop the well established CDAT/VCDAT software infrastructure.

Team

Coordinator of the effort:

- Valerio Pascucci, LLNL (Team Leader).
- Peer-Timo Bremer (LLNL): core visualization techniques, data comparison, topology.
- Marty Cole (Utah): deployment of visualization libraries.
- Jamison Daniel (ORNL): user interaction, case studies, testing of tools.
- Ming Jiang (LLNL): topological analysis, VTK components.
- Daniel Laney (LLNL): data management and streaming.
- Ajith Mascarenhas (LLNL): feature tracking.
- Claudio Silva (Utah): provenance-based comparative analysis.
- Xavier Tricoche (Utah): 2D topological visualization and analysis.

High-level Stakeholder Need(s):

- **Task A.** Deploy advanced visualization capabilities into the CDAT tool and create a clear path for similar integration in other tools.
- **Task B.** Extend the visualization software to incorporate domain specific requirements, data formats, and vector field visualization.

[26]http://www.cgd.ucar.edu/csm/
[27]http://www.earthsystemgrid.org/
[28]http://www.ipcc.ch/

- **Task C.** Support time-dependent and cross-dataset comparison, visualization and analysis.
- **Task D.** Develop new analytic capabilities for climate data (first deployed into CDAT and VCDAT).
- **Task E.** Integrate with VisTrails advanced framework for tracking and logging of internal state and provenance information.
- **Task F.** Develop a visualization and data analysis scenario for understanding of complex coupled phenomena such as the multi-scale dynamics the complete carbon cycle on earth.

Milestones/Deliverables/Objectives

Task A: Deploying Advanced Visualization Capabilities into Production Software Infrastructure.

The customers would like to integrate state of the art and new 3D visualization capabilities into the CDAT/VCDAT package for rendering of climate data. This need will be addressed by providing a Python integration for a visualization library based on a streaming data model, which addressed the main request of having a light weight and portable external package to be released with CDAT and easy to integrate with other deployment targets. The work will proceed in three steps:

- Defining a high-level interface into ViSUS.
- Creating a Python wrapper around the interface.
- Integrate the interface into both CDAT (the command-line tool) and VCDAT (the GUI version).

The final deliverable for this stage is a 2D/3D contouring and 3D volume rendering interface for (V)CDAT for regular (Cartesian) grid data.

Task B: Customized Extensions. Climate researchers use a variety of special data formats and have some domain specific requirements not common in generic scientific visualization tools. In particular, we plan to provide visualization with data projected onto a model of the earth (sphere or geoid), enhance the 3D visualization with underlying geospatial information (i.e. satellite images or boundaries of the nations), and to natively handle various types of meshes used in the climate community (e.g. geodesic grids, curvilinear grids, etc.). Furthermore, we plan to address a need for vector field visualization tools, for example, using streamlines and/or LIC based techniques. The deliverables include extending the visualization packages with the specialized techniques needed and the development of vector field visualization package.

Task C: Multi-dataset/time-dependent visualization and analysis. One important aspect of climate research is the constant need to compare and analyze data from different models or from the same model at different points in time. To support this need the visualization should be extended to incorporate multiple linked views, and interactive exploration of (potentially long) time series. The deliverables in this area include the computational infra-structure to handle large time-dependent data sets in (V)CDAT and a user-interface supporting cross-model and cross-time comparison and exploration.

Task D: Advanced analytics. The customers have a strong interest in developing new analytics capabilities and having them integrated into (V)CDAT. Of particular interest are novel techniques for defining, analyzing, and tracking features as well as new multi-scale data comparison metrics. To this end we will aim at integrating large scale statistical analysis as well as topology based techniques into CDAT and provide high level interfaces for their usage. Furthermore, we are

developing new data comparison metrics which will provide multi-scale structural metrics. There is also a need for new topological analysis based on robust combinatorial data processing extending recent work developed at LLNL for conservative vector fields. This includes finding critical point, creating the topological skeleton, and robustly simplifying vector field topology. The deliverables in this area include a suite of multi-scale analytical tools both vector and scalar data to be integrated into (V)CDAT.

Task E: VisTrails integration. Demonstrate with concrete examples how CDAT users can integrate the VisTrails infrastructure and provide the users with improved productivity. In particular, logging of web based operations, improved data and process provenance, new user interfaces for large parameter studies. We expect that the list of needs to evolve as evaluation of preliminary work progresses, and through collaboration with the CDAT development team.

Task F: Case studies based on coupling of multiple models and multi-scale dynamics. The ultimate goal of our work is to provide visualizations and analysis techniques that enable a new deeper understanding of the dynamics of climate models. We will include in our activities direct support for the analysis of particularly complex and interesting case studies. On probable target will be the study and understanding of the global carbon cycle, including atmospheric chemistry, integration of land and ocean ecological processes, and their coupling with climate. This requires data analysis and visualization incorporating several models to answer fundamental questions related to climate variability and global change at time scales ranging from decades to centuries. This is a major challenge requiring close collaboration with the climate researchers.

Subtasks Task A: Deploying Advanced Visualization Capabilities into Production Software Infrastructure.

Task A1. Define and wrap interface to contouring library (8 weeks). **(Completed.)**

- Define low-level interface: select the appropriate C-library functions (1 day).
- Create typemaps for SWIG to interface Pythons Numeric Arrays with C-data structures (1 week).
- Test typemaps and integrate robust type-checking and useful exceptions to speed-up later debugging (1 week).
- Create contouring Python module regression tests (1 week).
- Implement stand-alone C-based OpenGL routines for performance reasons and to avoid PyOpenGL problems (see risks) (1 week).
  - Flat shading.
  - Smooth shading.
  - Wireframe.
- Create stand alone Tkinter-based sandbox for testing and later integration (2 weeks).
- Regression testing the code by comparing with the native ViSUS package, performance check (1 week).
- Develop initial user-interface to steer contour visualization (trackball, iso-value changes, File IO etc.) (1 weeks)

Task A2. Define and wrap interface to a basic volume rendering library (9 weeks). **(Completed.)**

- Define low-level interface (1 day).
- Create typemaps (2.5 days).

- Test typemaps and implement type checking (2.5 days).
- Create volume rendering Python module (1 week).
- Simple regression tests (1 week).
- Create alternative compilation targets for use of 2D texture, 3D texture, and mesa software rendering (2 weeks).
- Implement C-based OpenGL routines and interface the texture based based rendering into PyOpenGL (1 week).
- Integrate the Python module into the sandbox (3 days).
- Regression test, validation, performance analysis (1 week).
- Extent the user interface of the sandbox to incorporate volume rendering (1 weeks).

Task A3. Non-orthogonal progressive data slicing: (9 weeks) **(in progress)**.

- Extract slicing library from the ViSUS data access module (2 weeks), completed.
- Define low level interface (1 day), completed.
- Create and test typemaps (2 days), completed.
- Create slicing Python module (1 week), completed.
- Regression testing (1 week), completed.
- Implement C-based OpenGL routines for fast rendering (2 days), completed.
- Integrate slicing module into the sandbox (3 days), completed.
- Regression testing performance analysis (1 week), completed.
- Implement prototypical .idx file-format conversion at load time (1 week).
- Performance analysis of idx conversion viability assessment (1 week).
- Prototypical progressive display and viability evaluation (1 week).
- Extend user interface of the sandbox to incorporate non-orthogonal slicing (1 week).

Task A4. Prototypical CDAT integration restricted to one particular target platform (2.5 weeks), completed.

- Integrate the sandbox into CDAT by hand on some selected machines (e.g. a presentation laptop) (1 week).
- Evaluate the user-interface with Dean's team decide on necessary changes and additions (2 days).
- Discuss the need/choice of colormap/transfer function editor with Dean and/or "standard" color maps / transfer functions (1 week).

Task A5. Interface design, integration and adaptation (3 weeks), completed.

Task A6. Incorporate/implement the chosen method of colormap/transfer function editor (3.5 weeks).

- Colormap editor (1 week).
- Transfer function editor (2 week).
- Develop a small library of preexisting transfer functions and allow to grow the database (3 days).
- Test integration of transfer function design with contour spectrum and other histograms (1 week).

Task A7. Final (V)CDAT integration (5 weeks).

- Collect all necessary libraries into the CDAT distribution (2 days).
- Adapting the build system / uncouple libraries (2 weeks).
- Compatibility testing for Windows, Linux, and MAC platforms (4 weeks).

Task A8. (V)CDAT – VisTrails integration.

- Get feedback from Dean Williams and his team.
- Explore API XML specification.
- Possible tighter integration between spreadsheet and the vcs.Canvas (currently file based):
  - Identification of dependencies: The path forward depends on feedback.
  - Subtasks: Marty Cole to work on the list of tasks. Other members of the VisTrails development team have been working on improving functionality to make his work easier, including easier installation on Linux, and other platforms, improved spreadsheet functionality, and other enhancements to the system to make it easier to integrate.
  - Estimated duration, start date, completion date: started work Nov 07, Duration and completion depend on feedback.

Risks.

- RISK: ViSUS code unreliable. Likelihood: low. Impact on project: high. Remediation: Debug/rewrite. However, the ViSUS libraries have been thoroughly tested and a mission-critical bug is unlikely. Team members are familiar with the code and eventual unexpected debugging will not bee too time consuming. Potentially, the volume rendering could be replaced with the SCIRun library (which might be a later target anyway due to its potential for higher image quality).
- 2. RISK: OpenGL/Python integration: The current OpenGL extension to TkInter called Togl is no longer supported and has some known compilation problems. Likelihood: low. Impact on project: medium. Remediation: We are expected to use Togl only in a very minimalistic way in which no problems have occurred in the past. Williams' team has extensive knowledge of python and seem well able to fix most foreseeable problems. The worst case scenario is to move toward a standalone GUI started from but not integrated into CDAT (i.e. a separate main loop).
- RISK: Moving to NumPy: At some point the CDAT project will move from Numeric.arrays to NumPy which will require a change of all wrapping code. Impact on project: high. Remediation: Depending on when this move happens, it will likely require at least one-two weeks to switch the code base. There is a small possibility that necessary functionality will be missing from NumPy and a more extensive re-write might be necessary. Looking at the current state of Numeric and NumPy this seems unlikely. To avoid major problems we will use only constructors that are already used in CDAT so that we can adopt fixes developed in the ESG group.
- RISK: Volume Rendering compatibility: The current implementation of the volume rendering library relies on some OpenGL features (3D texture mapping) for higher performance. These features may not be present on all necessary platforms. If that is the case we will need to provide alternative compilation targets that fold automatically back to the basic slower code based on 2D texture maps. Likelihood: medium-low. Impact on project: high. Remediation: Ignoring the performance issue the render can be extended to include a software-only version based on Mesa. Another option is to switch entirely to the SCIRun volume render.

- RISK: Transfer function design: The design of good transfer functions is non-trivial even for experts and it is unclear whether there exists a "good" editor for general users. Creating an editor might therefore be less a problem of time investment but rather of developing a novel interface paradigm. The transfer function is an integral part of any volume rendering and without it volume rendering is unlikely to be adopted in the climate community. Likelihood: medium. Impact on project: medium (with a critical impact on the volume rendering aspect). Remediation: There exists research on automated transfer function design that could be applied with some time investment. Furthermore, it might be possible to write a meta-editor which adapts pre-existing transfer functions to a limited set of problems. A large library of pre-defined transfer functions would also alleviate the problem. Finally, a well designed tutorial could be used to educate CDAT users in the use of one or multiple of the advanced transfer function editors developed within the visualization community (e.g. multi-dimensional transfer functions).
- RISK: CDAT integration: CDAT is required to build on a large variety of platforms some of which might be drastically different from "standard" visualization systems. Potentially there exist platforms which miss crucial components (e.g. OpenGL drivers) which cannot be included into the CDAT distribution. Likelihood: very low (we do not expect to be dependent on any advanced OpenGL feature and we will run the code on low end platforms such as a laptop). Impact: low (high impact only for a few users). Remediation: We will make sure that our code compiles on the systems of our direct costumers. Other users in the general community will need to turn off our module or upgrade their system.

Task B: Customized Extensions.

Task B1. Integrate geospatial information (14 weeks).

- Create file reader for the image formats used in CDAT (2 weeks).
- Wrap NGA code for high resolution imagery and XML data (2 weeks).
- Implement composite rendering superimposing contours/volume rendering over maps (3 weeks).

- Avoid storing multiple data copies.
- Automatic registration between 2D and 3D data.
- Integrate two glut main loops and fix any issue with interfering global rendering variables and structures.
- Integrate code into the sandbox (2 weeks).
- Regression testing (1 week).
- Porting code to selected CDAT platforms and integrate into CDAT build system (4 weeks).

Task B2. Integrate terrain visualization (9 weeks).

- Implement SOAR index mapping (2 weeks).
- Explore SOAR-based data re-ordering for performance [this is similar to the idx conversion under A3] (2 weeks).
- Integrate code into the sandbox (2 weeks).
- Regression testing (1 week).
- Porting code to CDAT platforms integrate into build system (2 weeks).

Task B3. Off-line rendering and Movie creation (17 weeks).

- Develop file formats and readers and writers to save and load scenes (XML parser for saving and loading view point, color map, interface configuration, etc.) (2 weeks).
- Incorporate off-line rendering into a virtual window of any arbitrary resolution (i.e. "Print" button) (4 weeks).
- Off-line rendering into compressed image formats (e.g. PNG) (1 week).
- Create movie pipeline for static view-points (2 weeks).
- Create movie pipeline for dynamic view-points (i.e. fly-through) (3 weeks).
- Parallelize movie creation for selected batch systems (2 weeks).
- Regression testing (3 weeks).

Task B4. General embedding of regular domains : Allow regular grids to be warped by a general XYZ mapping (17 weeks).

- Adapt iso-contouring code to warped grids (2 weeks).
- Regression testing (2 weeks).
- Adapt volume rendering to warped grids (8 weeks).
- Regression testing (3 weeks).

Task B5. Tetrahedral and general meshes (4-6 months).

- Decide which are the important grid types that should be supported (1 week).
- Create file readers (2 weeks).
- Implement re-sampling procedures to deal with grids that cannot easily be mapped to the visualization algorithms (4 weeks).
- Adapt visualization tools to the remaining grid types (2-4 months).

Task B6. Visualization of computational domains for debugging (6 weeks).

- Implement high-performance OpenGL routines to visualize the underlying grid structure (rather than the data) (1 week).
- Wireframe.
- Hidden lines.
- Imploded cells.
- Integrate into sandbox (3 days).
- Develop query tools and visual aids (3 weeks).
- Highlighting.
- Picking.
- Range queries.
- Clipping planes.
- Porting to CDAT platforms (2 weeks).

Task B7. Vector field visualization: (4-6 months).

- Implement and adapt standard vector field visualizations such as streamlines and LIC based visualization for 2D gridded data (1-2 months).
- 3D streamline visualization of regular grid data (1-2 month).
- Visualizing derived quantities such as vortices, shockwaves, etc (1-2 month).

- Integrate visualization code into the sand box (3 weeks).
- Regression testing (2 weeks).
- Porting to CDAT platforms (2 weeks).

Risks and Opportunities

- RISK: Volume rendering for warped and general meshes: Adapting the volume rendering to non-cartesian meshes is highly non-trivial and likely to be computational expensive since it requires sorting the data and error prone due to many special cases. Likelihood: high. Impact on project: medium (high for the volume rendering aspect). Remediation: We have some code available that implements a general software renderer for tetrahedral meshes. Other options are to use better code that might be available in Utah or regrid the data.
- RISK: Unusual grids: The climate community might use grids very different and non-adaptable to visualization. Likelihood: medium. Impact on project: low. Remediation: One can always perform a resampling on either regular or tetrahedral meshes.
- OPPORTUNITY: unstructured tetrahedral meshes may not be needed. Likelihood: unknown. Impact on project: low. Remediation: We will adjust our plan and achieve our targets sooner.
- RISK: User interface for vector fields: Vector field visualizations depend on many parameters such as streamline density and location, integration step size, etc.. It is not clear how these can be exposed in a meaningful way. Likelihood: medium. Impact on project: low. Remediation: More work on the interface and if necessary some trial and error session with the users to determine useful defaults.
- RISK: Performance problems dues to the extra cost of grid mapping andor multiple rendering threads: If used in full there will be up to 4 rendering threads running at any one time (iso-surfaces, volume rendering, terrain visualization, 2D imagery) which will impact performance. Likelihood: medium. Impact on project: medium. Remediation: Depending on the severity more time can be allocated to optimize the code andor make better use of modern GPU's if available. Features can always be selectively turned on and off which should allow useful exploration in any case. For some users a parallel rendering framework would be a option andor remote rendering.

Task C: Multi-data and settime-dependent visualization and analysis.

Task C1. Multiple linked views (7 weeks).

- Extend the sandbox interface to allow multiple linked viewports showing multiple data sets (4 weeks).
- Support multiple data sets in memory at the same time.
- Cross data set registration.
- Allow dual and multiple use of interface functionality, e.g. use the same iso-surface slider for all viewports.
- Link and unlink visualization parameters such as iso-value, viewpoint, transfer function, etc.
- Port to CDAT platforms (3 weeks).

Task C2. Support time-dependent visualization (2-4 months).

- Develop data management framework to load and cache time steps for interactive visualization (1-2 months).
- Explore the feasibility of converting the data to visualization-friendly data formats ,i.e idx, (1-2 months).

Task C3. Parallel and Remote visualization (4-7 months).

- Spread the visualization tasks over multiple threads to utilize multi-core machines (1-2 months).
- Distribute the visualization across a local network (2-3 months).
- Enable remote streaming visualization utilizing a client-server architecture (1-2 months).

Risks.

- RISK: Code compatibility: Cross-platform parallel code is very difficult to develop and might be too ambitious. Likelihood: high. Impact on project: low (most users are not expected to have easy access to parallel resources). Remediation: Concentrate on few common platforms in particular those at visualization centers (LBL, LLNL, ORNL) and thread based implementations. Jump directly to remote visualization using shared resources.
- RISK: Data size: Depending on the common data sizes single desktop, interactive, time-dependent visualization might simply not be feasible. Likelihood: medium. Impact on project: low. Remediation: Remote visualization and a bigger focus on hierarchical techniques.
- RISK: Remote vis: there may be problems with firewalls when dealing with remote data streaming. Likelihood: high. Impact on project: medium. Remediation: Worst case: remote visualization will be enabled only within selected networks. Expected solution: use of SSH tunnels for all “difficult” situations.

Task D: Advanced analytics. Advanced analytics (23-25 month + vector field research).

Task D1. Integrate 2D structural analysis into CDAT (3 months).

- Extract a library from topology applications (3 weeks).
- Wrap the low-level library (2 weeks).
- Integrate into sandbox (2 weeks).
- Topology interface (critical points, manifolds, cells).
- Topology based highlighting, picking.
- Regression testing (2 weeks).
- Port to CDAT platforms (2 weeks).
- Interface evaluation (1 week).
- Create high-level command-line interface for CDAT (2 weeks).
- Create high-level user-interface for VCDAT (1 week).

Task D2. Flexible feature definitions: Create an interface allowing the user to create feature definitions on the fly (5 weeks).

- Interface definition (2 weeks).
- Implementation (3 weeks).

Task D3. Robust feature tracking (7 months).

- Integrate merge tree code with flexible feature definitions (4 weeks).
- Develop hierarchical feature tracking (3 months) (combustion project).
- Specialize topological tracking to climate related problems (1 months).
- Porting to CDAT platforms (4 weeks).
- Interface and visualization (1 month).

Task D4. Large scale statistics (6-8 months).

- Develop fast hierarchical methods to access data and collect statistics for large data (2-3 months).
- Create portable library and Python wrapper (4-6 weeks).
- Integrate into sandbox (2-4 weeks).
- Evaluate practical usability and impact (1 week).
- Specialize tools and interface for climate analysis and adapt the interface (1-2 months).
- Porting to CDAT platforms integrate build system (1 month).

Task D5. Multi-scale data comparison (6-10 months).

- Develop mathematical framework (1-2 month).
- Extend code base (e.g. topology code, statistics library) (1-2 month).
- Develop and evaluate different metrics (2-3 month).
- Interface and visualization (2-4 weeks).
- User evaluation (1 week).
- Porting to CDAT platforms integrating build system (2-4 weeks).
- Specialize interface visualization for climate analysis (4-6 weeks).

Task D6. Vector field topology (15-24 month).

- Robust definition of critical points (1 month).
- Developing "discrete" 2D vector field theory (4-5 month).
- Robust computation of 2D vector field topology (2-3 month).
- Defining topological simplification, error metrics, and hierarchy (2-4 month).
- Extension to 3D vector field topology (6-11 month).

Risks and Opportunities.

- Opportunity in stage D3 we may be able to unify and leverage the work in support of combustion simulations. Likelihood: high. Impact on project: high. Remediation: adjust the plan and deliver the tools earlier than planned.
- RISK: Research success uncertain: In general, many of the techniques depend on additional fundamental research that might not pan out or take significantly longer. Likelihood: unknown. Impact on project: medium. Remediation: Fall-back to traditional techniques.
- RISK: Flexible feature definition: A general framework that allows random data access during the feature definition can become a severe performance problem. Likelihood: medium. Impact on project: low. Remediation: Restrict the feature definitions to filters of certain width, streaming filters, and/or integrate progressive framework.

- RISK: Feature tracking: Features in climate models might be too complex to be defined in closed mathematical terms. Likelihood: low. Impact on project: medium. Remediation: Multi-layered feature definition on derived quantities. For example, allow features to be defined using a combination of multiple input fields in combination with derived quantities (vorticity etc.).
- RISK: Large scale statistics: this combination with large scale data access and statistical methods is new and will require close collaboration of the LLNL team with the ORNL team on the development of a working prototype. The novelty increases the risk of delaying the deployment. Likelihood: medium. Impact on project: low. Remediation: CDAT already carries an interface to the R statistics package from which one can reuse or adapt components if specialized techniques fail.
- RISK: Multi-scale data comparison: The ideas underlying this aspect of the plan have not yet been tested and probably require some more research. Likelihood: medium. Impact on project: low. Remediation: The exists a large array of standard comparison techniques that can be implemented if the novel specialized metrics fail.
- RISK: Vector field topology: The theory for robust/combinatorial vector field topology has not been developed yet. Experience in the scalar case suggests that there will be a significant ramp-up time but also great benefits. Likelihood: unknown. Impact on project: low (vector field topology is only a small specialized aspect). Remediation: Several VACET members have experience in numerical methods for computation of Vector field topology (especially Xavier Tricoche and Ming Jiang) and their results could be used as an intermediate/substitute solution.

Task E: VisTrails integration. No data at this time.

Task F: Case studies based on coupling of multiple models and multi-scale dynamics.

Task E1. Work with the users to develop a high-impact case study based only on visualization tools (2 months).

Task E2. Introduce analytical component and quantitative feedback complementing the visualization (2 months).

Task E3. Refine and integrate the visualization and analytics components, develop conclusions and document results with images, videos and writing a technical report (4-6 months).

Accomplishments

- We have developed an initial python wrapper around the ViSUS code base. This result enables direct access to the ViSUS scene graph and its full manipulation via python scripting. In the process we have made available fundamental features such as isosurface computation, volume rendering and computation of arbitrary slices while developed new nodes that add high quality labels to the scene (using an external font library) and tick marks for axes and color bars. The image below (Figure 7 has been generated by the python script `Cloudiness.py`. In the current interaction with Deans group we want to refine the API exposed in python since the current version is simply exposing all parameters available in the C++ implementation.
- We are exploring alternative scene layouts for data exploration and movie creation. For example the image below (Figure 8 shows a possible layout including both embedded 3D data and slices on the side. This type of layout can increase both the realistic presentation of the information and the effective communication scientific the results. After completing

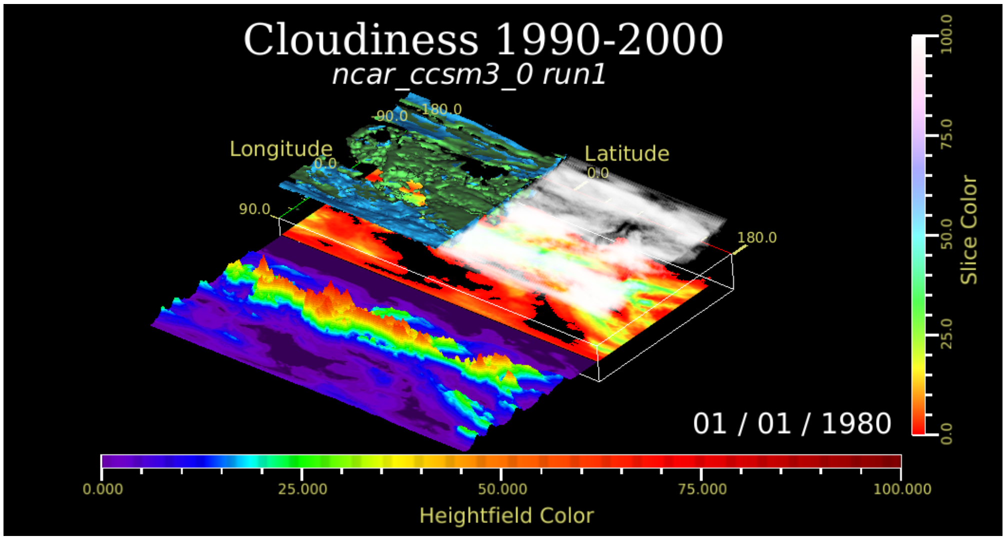

Figure 7: This image represents a new capability – a prototype integration of advanced visualization capabilities into the ESG-sponsored Climate Data Analysis Toolkit. This result proves the feasibility of the VACET approach to layering new capabilities onto high-quality software for use by a broad community.

our discussion with John Drake, Jamison Daniel, and the rest of the climate team at ORNL we will improve the layout and content of the scene and generate a complete animation via scripting embeddable in CDAT. This will allow for example to give the script to the users and produce similar presentation by simply changing the data source in the python script.

- VisTrails integration
  - We created installation scripts/instructions for Linux and Mac. These install CDAT and VisTrails into a shared python installation. This makes it easier for collaboration to take place.
  - Created XML format for automatic wrapping of target API's. This file can be read to generate a VisTrails package `__init__.py`, which is the main wrapping file to load a package into VisTrails. The hope here is to generate the API description for selected API from CDAT, and quickly create the package specification from it. This way as API changes over time, the package can be quickly and automatically updated. The main question with this strategy, is how hard it is to generate the API descriptions for target functionality once it is identified.
  - It may become important to write a wrapper for the main CDAT viewer, such that it is usable directly inside the VisTrails spreadsheet. We wrote such a wrapper using the SCIRun viewer as a testbed. The experience gained from this positions us well for doing similar work for CDAT's main viewer. If the stakeholder and his team find it important, we should be able to collaborate a similar solution for CDAT.

Work Targets for Next Three Months

- Restructure the Python to develop a simpler API.
- Deploy new version of the ViSUS scene graph.
- Improve the stability of the VCDAT integration.

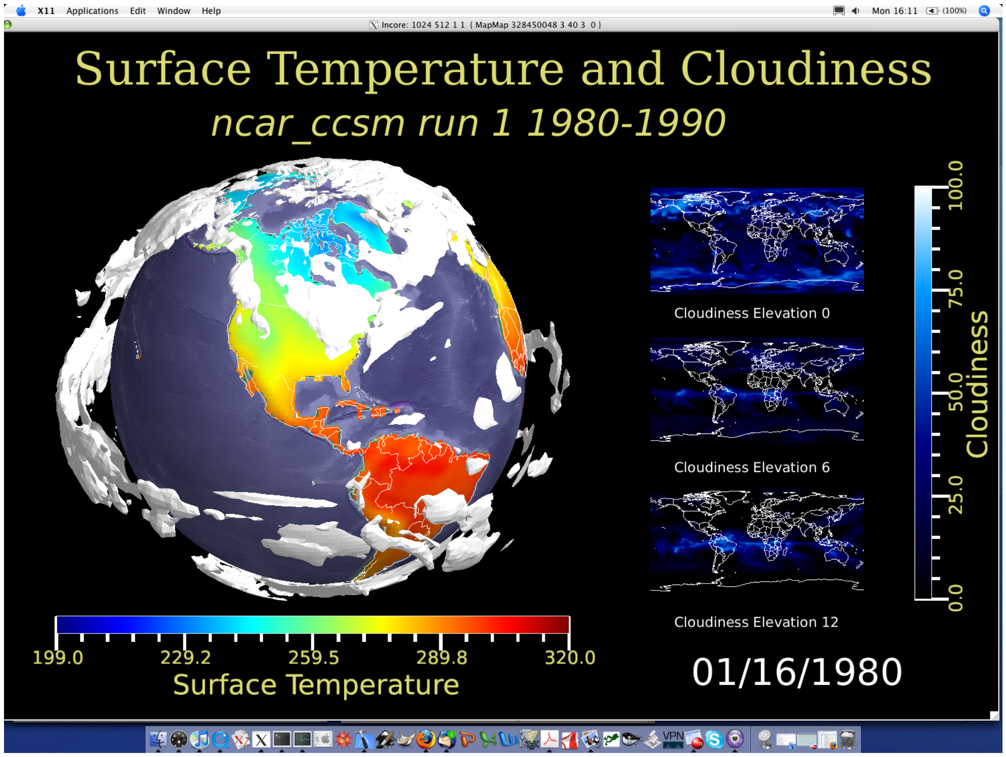

Figure 8: This image shows use of both traditional 2D and new 3D visualization capabilities within CDAT. This type of layout can increase both the realistic presentation of the information and the effective communication scientific the results.

Work Targets for Next Six Months

- Deploy first version of ViSUS integrated in VCDAT.
- Improve/simplify user interface.
- VisTrails.
  - Get feedback from Dean and his team.
  - Identify best path forward to wrap CDAT interface. E.g., XML specifications?.
  - Determine if a tighter integration between the canvas and the spreadsheet is a priority or not.
  - Identify a target to show parameter exploration capabilities.
  - Support any CDAT users who wish to integrate their scripts.

Work Targets for Next Twelve Months

- Deployment of basic 3D visualization package in (V)CDAT.
- Determination of the most desired extensions (vector fields, multi-data set, time-dependent) and initial results.
- Research prototype for extended analytic capabilities.

12.4 Combustion

The VACET team is performing R&D on several projects aimed at providing new visual data understanding capabilities to the combustion research community.

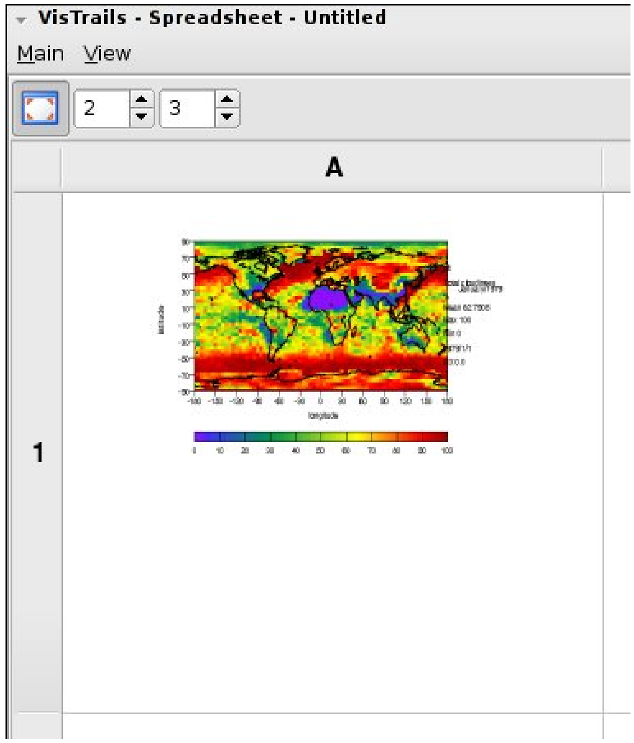

Figure 9: Resulting images viewed inside he VisTrails spreadsheet.

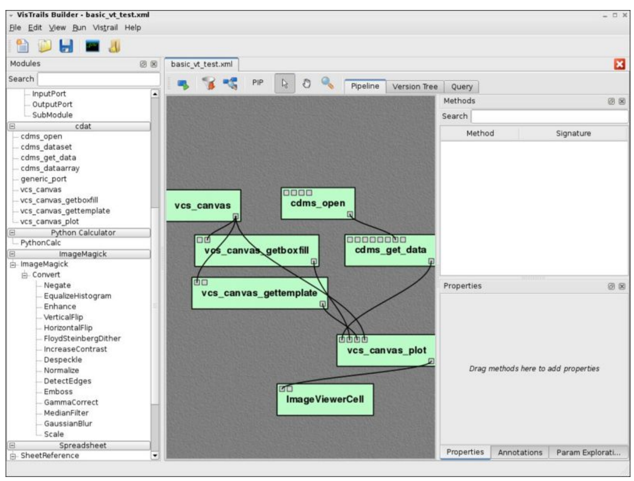

Figure 10: The pipeline of wrapped CDAT interface modules that produced the image in the spreadsheet.

12.4.1 Advanced AMR-Based Temporal Analysis and Visualization

VACET Team Members: Hank Childs, LLNL (Team Leader); Gunther Weber, LBNL.

Customer/stakeholder: Marc Day, LBNL, of the CCSE group under John Bell.

Stakeholder need(s).

In this case, Marc would like to be able to identify a surface of interest (a contour that represents the start of the flame) and be able to walk backwards in time using a vector (e.g. use the velocity to determine where the chemicals currently on the flame surface were located at the previous time slice), and then repeat all the way back to the beginning of the simulation. As the surface is traced backwards (to create a volume), Marc would like different scalar variables (such as species information) to be evaluated at each time. Once these volume are created, Marc would additionally like to do analysis on these volumes, by calculating ratios using combinations of surface areas on the surface, volumes of the swept out prisms, and the scalar fields on the mesh.

Subtasks/milestones/deliverables.

Task 1. Create a VisIt operator that will start with an initial surface and iterate backwards through time, generating the swept volume as it goes.

Task 2.

Establish a file format that can capture this volume without exceeding available memory. That will likely involve dumping the results from each time slice as it goes.

Task 3. Create a VisIt operator that will modify the contour to no longer contain extremely small triangles. These triangles create numerical problems. “Small” is measured with surface area and high aspect triangles must also be removed.

Task 4. Add routines to analyze the resulting files and create the 1D curves Marc is truly interested in.

Task Dependencies Dependencies: Tasks 1 and 2 will be performed concurrently. Task 3 is independent of 1 and 2. Task 4 depends on 1, 2, and 3.

Estimated duration: Tasks 1 & 2: 1 week. Task 3: 1 day. Task 4: 1 week.

Accomplishments

Tasks A1 and A2 are completed.

Work targets for the next six months.

Complete Tasks 3 and 4. Hank Childs will be doing the work with Weber observing and providing support.

Risks

1. VisIt infrastructure for parallel point location is buggy. Likelihood: low. Impact on project: medium. Remediation: would have to debug point location routines in VisIt (this could take a while). Affected: Task 1.
2. Creating “acceptable” triangle mesh contains hidden costs. Likelihood: low. Impact on project: high. Remediation: should iterate with Marc to better understand what will and will not work. Marc suggested using QSlim, but that is a heavyweight project (although it caps the cost). Affected: Task 3.
3. Analysis of resulting volume may go beyond the ratios that Marc initially mentioned. Likelihood: high. Impact on project: medium. Remediation: I don’t think there is a remediation here. The goal is not to “get in and get out”. It is to help Marc solve science problems. If this requires more development than initially foreseen, then that’s what it will take. However, the likelihood that this becomes an extremely large time investment (i.e. multiple man-months) is extremely low. Affected: Task A4.

Timeline Next Six Months

Start date: 4/6/07. Targeted end date: 12/31/2007.

Next 12-24 months. This project will complete in the next six months.

12.4.2 Topological Combustion Data Analysis – Part I

VACET Team.

- Valerio Pascucci, LLNL (Team Leader).
- Peer-Timo Bremer (LLNL) topological feature extraction.
- Daniel Laney (LLNL) data management and streaming.
- Ajith Mascarenhas (LLNL) feature tracking.

Customer/Stakeholder(s). Jacqueline Chen, SNL.

Customer/Stakeholder need(s). Topological characterization of features in a combustion process such as extinction pockets. Development of a robust tracking scheme and practical use for reliable quantitative analysis. These capabilities offer traction on the problem of visual and quantitative analysis, something very much needed as part of the scientific data understanding process for combustion research. These capabilities are not provided by any other source, particularly not in scalable visualization tools (production or research).

Deliverables to meet Stakeholder Needs. Deliver topological analysis software to aid understanding of combustion process. Jacqueline Chen's group works with large 2D and 3D time-varying combustion simulation datasets, and needs analysis tools to aid their understanding of the combustion process. They are interested in analyzing the data to detect the dominant re-ignition mechanism of extinction pockets.

Task 1. Identify data formats, develop data read/write modules, format interchange modules, assess time to marshal data (2-3 weeks) **(in progress).**

- Develop read/write software from Fortran Binary / NetCDF to Brick of Value (BOV) for VisIt and IDX for ViSUS (1 week per format).
- Test software. (3 days per format).
- Data movement and preprocessing
    1. Expected data size and processing steps to prepare data for analysis: (1) current "Excite Jet" 3D data set is 43 GB per time step (242time steps), stored as a single file per time step on HPSS tape; (2) the expected preprocessing pipeline will be: (a) Download one timestep from tape to disk (either NFS or Lustre): 1 hour minimum (assuming available disk space); (b) extract required fields into the proper format for our tool chain 15 minutes per field per time step; (c) Store these files back to tape and transfer to LLNL to mitigate risk 1 minutes (local HPSS), 90 minutes LLNL
    2. Estimated time to preprocess the "Excite Jet" data set: (a) 250 hours (10 days, extreme minimum) to download files from tape; (b) 60 hours (3 days, extreme minimum) to extract 1 field from tape, interleaved with previous bullet; (c) 370 hours (16 days, extreme minimum, in parallel with above) to move extracted fields to LLNL filesystem.
    3. Quality assurance of data by performing sanity check of value ranges. (E.g. all pressure values must be within acceptable ranges.) (3 days.)

Task 2. Develop basic visualization software for preliminary study and visual assessment of the preprocessed data. (4 weeks.) **(completed – Fall 2007).**

Task 3. Develop feature segmentation and tracking software for time-varying 2D combustion data. (30 weeks.) **(completed in 19 weeks; Fall 2007).**

Feature Segmentation

- Develop 2D feature segmentation software using 2D Morse-Smale complexes. (3 weeks.)
- Test 2D feature segmentation. (2 weeks.)
- Present preliminary results to analyze a representative small data-set to stakeholder for feedback. (1 day).
- Iterate with Incorporate changes and add functionality as per feedback. (4 weeks.)

Feature tracking

- Develop algorithm to track selected features over time. (2 weeks.)
- Develop library to unify geometric ID's across pre-processing, feature extraction, and tracking tools (1 week)
- Implement preliminary version of algorithm. (4 weeks.)
- Test tracking algorithm. (2 weeks.)
- Write proof-of-concept software to visualize and interact with segmentation and tracking data. (3 weeks.)
- Present visualization and tracking software to stakeholder for feedback.
- Develop software into full-fledged tool with controls for:
  - Refining segmentation. (1 week.)
  - Sensitivity of tracking. (1 week.)
  - Simplifying tracking data. (1 week.)
  - Presentation at various scales. (1 week.)
- Test software: correctness, robustness, memory usage. (2 weeks.)
- Present software to stakeholder and get feedback. Iterate if necessary. (1 day.)

Task 4. Develop feature segmentation and tracking software for full-scale 3D combustion data. (Total: 70 weeks.) **(in progress).**

The design of the software developed in this stage will rely heavily on the lessons learned in stage A3 and will be crafted to integrate it into the visualization and interaction software developed for the 2D case. Special emphasis will be placed on handling large data-sizes.

Feature Segmentation

- Develop strategy to handle increase in data-size due to increase in dimension.
  - Develop streaming algorithm for 3D Morse-Smale complex construction. (4 weeks.) (completed)
  - Develop parallel algorithm for 3D Morse-Smale complex construction. (4 weeks.)
- Implement streaming algorithm for 3D Morse-Smale complex construction. (8 weeks.) (completed)
- Test streaming algorithm for 3D Morse-Smale complex construction. (8 weeks.)
- Implement parallel algorithm for 3D Morse-Smale complex construction. (8 weeks.)
- Test parallel algorithm for 3D Morse-Smale complex construction. (8 weeks.)
- Present results and obtain feedback from stakeholder. (1 day.)
- Choose best of streaming and parallel strategies and refine software and develop to deployment stage. (4 weeks.)

Feature Tracking

- Develop strategy to handle increase in data-size due to increase in dimension.
  - Develop streaming algorithm for feature tracking. (3 weeks.)
  - Develop parallel algorithm for feature tracking. (3 weeks.)
- Implement streaming algorithm for feature tracking. (4 weeks.)
- Test streaming algorithm for feature tracking. (4 weeks.)
- Implement parallel algorithm for feature tracking. (4 weeks.)
- Test parallel algorithm for feature tracking. (4 weeks.)
- Present results and obtain feedback from stakeholder. (1 day.)
- Choose best of streaming and parallel strategies and refine software and develop to deployment stage.
  - Integrate 3D segmentation and tracking into visualization and interaction software already developed in stage A4. (4 weeks.)
  - Present software to stakeholder and get feedback. Iterate if necessary. (1 day.)

Task 5. Develop critical point tracking software using Jacobi curves for time-varying 2D data. (Total: 15 weeks. Completed: 7 weeks.)

This component complements the results in Task 4 and provides the necessary degree of robustness.

- Identify appropriate interpolation functions based on data-format. (3 days.)
- As a starting point, evaluate and choose simplest interpolation and develop/adapt algorithm to compute Jacobi curves. (2 weeks.)
- Implement algorithm. (3 weeks.)
- Test algorithm on representative data-set. (2 weeks.)
- Develop visualization and interaction tool. (With proper design we should re-use software of A3) (1 week.)
- Present preliminary results to stakeholder and get feedback (1 day.)
- Develop software to deployment stage based on feedback. We incorporate other interpolation schemes at this stage. (4 weeks.)
- Stress test with available data-sets. (2 weeks.)

Task 6. Develop Jacobi curve computation software for time-varying 3D combustion data. (Total: 20 weeks. Completed: 3 weeks.)

As in the previous stages, going from 2D to 3D requires careful analysis of the impact of data-size increase and strategies to factor in this increase.

- Identify impact of data-size increase on Jacobi curves computation algorithm. (3 days.)
- Develop streaming algorithm to compute Jacobi curves using a small memory footprint. (4 weeks.) (Completed 3 weeks of effort.)
  - Develop test for determining if an edge in the underlying mesh can be finalized.
  - Develop algorithm to classify criticality of finalized edges and to free its memory.
- Implement streaming algorithm. (4 weeks.)
- Test algorithm using representative data-set. (4 weeks.)
- Develop visualization and interaction software. (2 weeks.)
- Present results to stakeholder and obtain feedback. (1 day.)
- Incorporate feedback/new features and iterate (4 weeks.)

- Stress test with full-sized data-set and iterate if needed (2 weeks.)

Task 7. Develop theory for Jacobi curve simplification. (Total: 1-2 years)

This stage is a major undertaking including new important features not available in any tool. Success in this area will have an impact on several topological analysis techniques that use Jacobi curves and lead to software that can be applied in solving real-world science problems that are currently infeasible. We have already a good preliminary understanding of how to approach this problem.

1. Develop definitions of problem in restricted setting of time-varying functions.
2. Understand how cancellation of critical point pairs changes over time.
3. Develop understanding of how change in pairings modifies Jacobi curve structure.
4. Develop algorithm to compute when and how pairing changes occurs.
5. Develop a multi-resolution hierarchy of Jacobi curves based on structural simplification induced by critical point cancellations.
6. Extend theory and algorithms to general setting of pairs of functions on manifolds. In other words, redefine problem in step 1 after removing restrictions on functions and iterate steps 2 through 5.

Task 8. High quality visualizations and movies for publication and distribution.

1. Iterate with scientists to produce high quality visualizations for publication as figures (2 weeks)
2. Plan movies using low quality rendering (2 weeks)
3. Render frames for movies using Povray or other high quality rendering package 50-200 hours per movie.

Task Dependencies

- Stages 2 through 6 depend on 1, but can proceed after development of a read/write module for a representative data format.
- Although 3 and 4 can be performed concurrently, the lessons learned during 3 can speed-up development in 4. Therefore we set 3 to precede 4.
- Similarly 5 and 6 can be performed concurrently, but we set 5 to precede 6.

Risks and Opportunities

- **Transition from research prototype code to robust deployment code may be difficult.** Likelihood: high. Impact on project: low. Remediation: We introduce several testing and feedback steps in each stage to ensure that the stakeholder is satisfied with the software. Moreover, all new development effort will use the best software engineering practices and focus on designing modular software. Affected tasks: 2 through 6.
- **Resource scalability: Large resource requirements to handle extremely large datasizes.** Likelihood: high. Impact on project: moderate. Remediation: We have access to several large high-performance computers in LLNL, LBL, and ORNL that can handle such data sizes. Affected tasks: 4 and 6.

- **ORNL Lustre filesystem.** While ORNL will be standing up a center-wide production Lustre filesystem with a capacity of 1.6 petabytes within the next 5 months, this is a relatively new development and is likely to suffer initial roll-out difficulties. NFS at ORNL is slower and smaller, requiring manual intervention and more data movement to and from tape as analysis progresses. Likelihood: high. Impact on project: medium, we will plan for these issues. Remediation: moving preprocessed data to LLNL will enable us to spread computation among 6+ file systems and 4+ clusters to mitigate filesystem downtime.
- **Software does not scale appropriate to increase in data-sizes.** Likelihood: high. Impact on project: high. Remediation: We incorporate scalability as a design requirement in each of our tasks. In particular, our plan includes a design of visualization, feature segmentation, and Jacobi curve computation software to handle large data-sizes when moving from 2D to 3D. We will design fast memory-efficient streaming algorithms and also design parallel algorithms to compute 3D Morse-Smale Complex and Jacobi curves. Affected tasks: 4 and 6.
- **Change in requirements from stakeholder.** Likelihood: medium. Impact on project: moderate. Remediation: We add feedback steps in each stage to identify any change in requirements as soon as possible. We keep our software designs sufficiently general and flexible to quickly add more functionality as and when required. Affected tasks: 3 through 6.
- **Jacobi curves are too noisy.** Likelihood: high. Impact on project: moderate. Remediation: We have a practical simplification algorithm that removes noise and produces a coarse-to-fine hierarchy of Jacobi curves. Preliminary studies indicate that this algorithm has potential that we will develop. Stage A7 plans for a full-fledged investigation into the mathematics and algorithmic development based on a thorough understanding of the problem. Affected tasks: 5 and 6.

Accomplishments

- We had several meeting to discuss preliminary results and refine the targets of the topological analysis.
- Combustion data stored in ViSUS IDX format.
- Customized ViSUS viewer and installed at the Sandia cluster used by Jackie Chen's group. The image below 11 shows the simplified user interface developed for combustion data.
- Developed a function-on-surface spectrum to allow the evaluation of the scalar dissipation elements conditional on the stoichiometric mixture fraction. In Figure 12, the spectrum on the top allows selection the best selection of dissipation elements base on length and and area contained in a level set on the stoichiometric mixture fraction isosurface.
- Tested the streaming computation of merge trees for hierarchical segmentation of the combustion data into regions of interest.
- Integration of the feature tracking code with the merge tree computation.
- Developed a user interface for interactive presentation of the 3D segmentation information. This allows the user to select the level of refinement. The images in Figure 13 show tracking information for a given level of resolution(left) and a simplified version of the trajectories and segmented regions (right).
- We have collected shape statistics for a variety of segmentation levels. Figure 14 shows an extinction region and basic shape information from principal component analysis. In general we have collected distributions of shape information dependent on the level of resolution of the segmentation. The results are currently being evaluated by Jackie Chen.

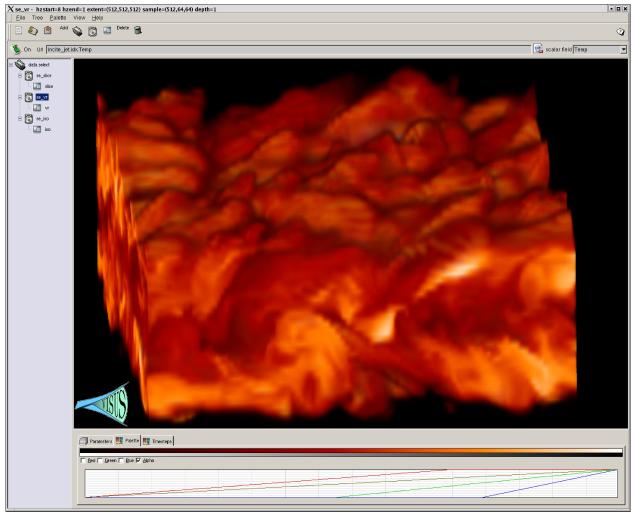

Figure 11: Preliminary test of multiresolution framework applied to a combustion simulation dataset.

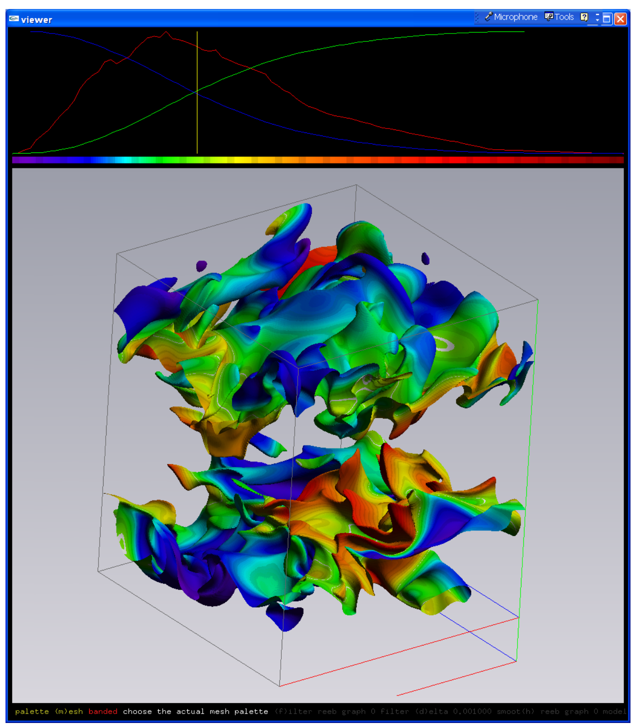

Figure 12: This image shows a new "function-on-surface spectrum" that allows the evaluation of the scalar dissipation elements conditional on the stoichiometric mixture fraction.

- Presentation of the paper: "On-line computation of Reeb-graphs: simplicity and speed" at the SIGGRAPH conference.
- Janine Bennett, Valerio Pascucci, and Ken Joy, Genus Oblivious Cross Parameterization: Robust Topological Management of Intersurface Maps. Proceedings of Pacific Graphics 2007.

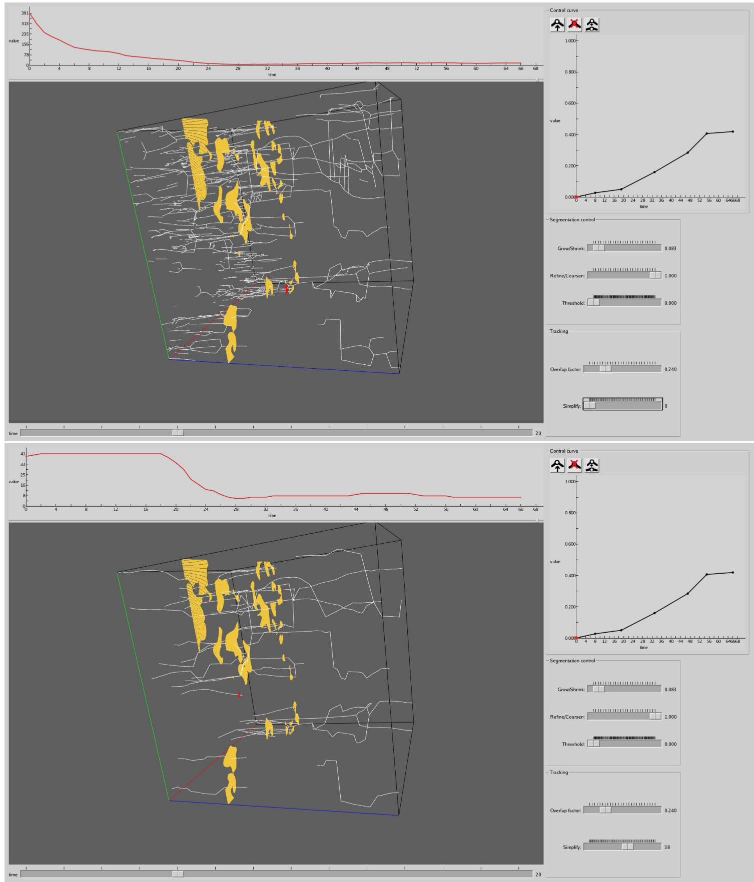

Figure 13: These images show a new prototype user interface for interactive presentation of the 3D segmentation information/results. Here, the user can select from among different multiresolution levels of refinement along with presentation of trajectories and segmented regions.

- Gunther Weber , P.-T. Bremer, and Valerio Pascucci, Topological Landscapes: A Terrain Metaphor for Scientific Data, IEEE Transactions on Visualization and Computer Graphics to appear, Proceedings of VIS 2007.

Work targets for the next six months.

- Tasks 1 and 2: In progress. Start date: 2/19/07.
- Task 3: In progress. Preliminary version of segmentation and feature tracking software already developed, and will be refined. Expected remaining duration: 11 weeks.
- Task 5: We will build on an existing code-base for Jacobi curve computation for time-varying 2D data-sets using two different interpolation functions. We have finished about 7 weeks of work, and will re-start on 05/15/07. Expected remaining duration: 8 weeks.

Within three months time, finish A3 and demonstrate feature extraction and tracking for time-varying 2D data to stakeholder. Within six months time, finish A5 and demonstrate Jacobi curve computation for time-varying 2D data to stakeholder.

Next 12-24 months.

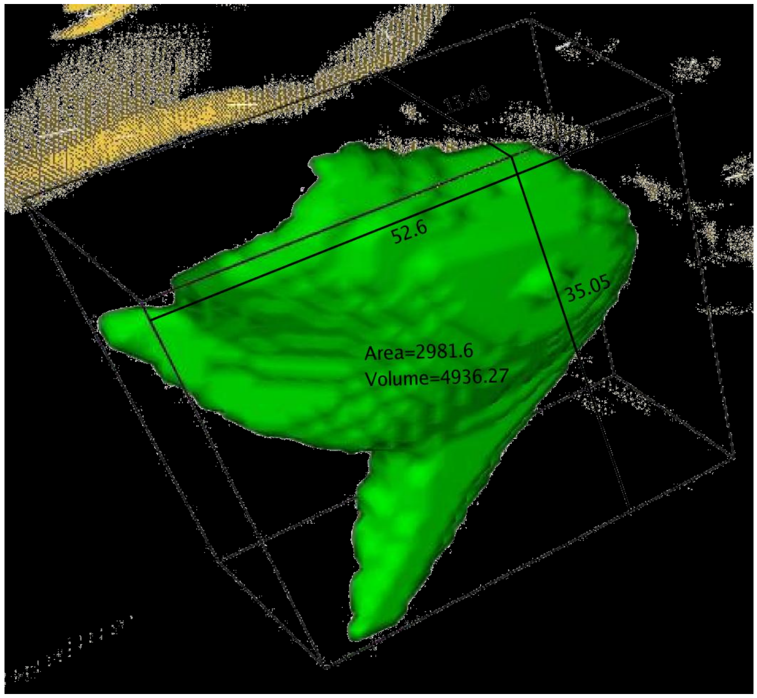

Figure 14: This image shows the results of combining analysis and visualization – we have computed and display feature characteristics, which in turn aid in accelerating scientific knowledge discovery.

- Task 4: Start date: 9/16/07. Expected duration: 70 weeks.
- Task 6: Start date: 6/01/07. Expected duration: 20 weeks.
- Task 7: Start date: 1/01/08. Expected duration: 1-2 years.

12.4.3 AMR Combustion Topological Data Analysis

VACET Team.

- Valerio Pascucci, LLNL (Team Leader).
- Peer-Timo Bremer (LLNL) topological feature extraction.
- Gunther Weber (LBNL) data management and topology-based analysis.

Customer/stakeholder: John Bell and Marc Day, LBNL.

Stakeholder need(s).

Topological analysis software to aid understanding of combustion process. John Bell's group works with large AMR 3D time-varying combustion simulation datasets, and need analysis tools to aid their understanding. They are interested in formation and evolution of extinction zones. They may be able to provide us with a boolean volume indicating regions that are "burning" or "non-burning," then we will perform a detailed analysis of the structural characteristics of the network of low temperature regions. They are also interested in using topological methods to identify those zones with topology-based methods. To this end we will use topological tools for both feature characterization and robust time tracking.

Subtasks/milestones/deliverables.

Task 1. Identify data formats, develop data read/write modules, format interchange modules, assess time to marshal data. (Total: 3-6 weeks.)

- Adapt custom file format readers (Marc Day's isosurface mesh with variable format, BoxLib) format. Marc Day's surface format is very similar to Timo's format, but a binary format. He has provided us with example code that should enable us to read his files quickly. He will also supply us with volume data in BoxLib format. These files will initially only contain a single patch at one resolution (size approx. 256 cubed). A simple example will show us how to get to a pointer to the actual data array. VisIt can read BoxLib files, so we can convert, view, analyze them using VisIt as well.
- Data movement:
  - We have computation time on davinci.nersc.gov.
  - CCSE code is known to work on davinci.nersc.gov.
  - We will perform work on davinci.nersc.gov.
- Initial data files:
  - Isosurface of temperature (isotherm) with one associate variable (consumption rate).
  - We do not have yet, but will obtain a 256^3 cubed volume file with two variables (fuel consumption rate, temperature) to start working with 3D methods and compare them to 2D surface based methods.

Task A2. Perform sample analysis on a 2D surface file (2 weeks).

Use Timo's code to extract extinction regions from family of isotherms. Determine, how dependent results are on temperature. If results are "reasonably stable" we may get away with looking at the 2D case. Present initial results to Day and Bell and solicit input/opinion.

Task A3. Perform sample analysis on 3D data.

- One quantity of interest is looking at a 3D sequence of fuel extinction and track the number of connected components for one isovalue or a set of isovalues.
- Apply Attila's existing code (Vis 2007) or code under development to these data sets.

Task A4. Develop topology extraction software for a small subset of the data. (36 weeks) **(in progress)**.

This stage will be used to understand the issues involved and for learning lessons before attacking the larger problem of analyzing the entire combustion dataset.

Feature segmentation

- Develop 3D Morse-Smale complex code to segment data. (6 weeks.) (3 weeks.) Completed.
- Test segmentation code on subset of dataset. (4 weeks.)
- Develop code to extract ridge-lines from segmentation (4 weeks.) (2 weeks.) Completed.
- Test extraction code on dataset. (2 weeks)
- Present preliminary results to stakeholders for feedback. (1 day.)
- Incorporate changes and add functionality as per feedback. (4 weeks.)
- Repeat step 4 based on feedback until the software is developed to required state.

Feature tracking

- Develop algorithm to track selected features over time. (2 weeks.)

- Implement preliminary version of algorithm. (3 weeks.)
- Test tracking algorithm. (2 weeks.)
- Write proof-of-concept software to visualize and interact with segmentation and tracking data. (3 weeks.) (1 week.) Completed
- Present visualization and tracking software to stakeholders for feedback. (1 day.)
- Develop software into full-fledged tool with controls for:
  - Refining segmentation. (1 week.)
  - Sensitivity of tracking. (1 week.) (Done.)
  - De-noising using persistence filtration. (1 week.)
  - Simplifying tracking data. (1 week.) (Done.)
- Test software: correctness, robustness, memory usage. (2 weeks.)
- Present software to stakeholder and get feedback. (1 day.)

Task A5. Develop feature segmentation and tracking software for full-scale 3D combustion data.

Feature segmentation

- Develop strategy to handle increase in data-size due to increase in dimension.
  - Develop streaming algorithm for 3D Morse-Smale complex construction. (4 weeks.)
  - Develop parallel algorithm for 3D Morse-Smale complex construction. (4 weeks.)
- Implement streaming algorithm for 3D Morse-Smale complex construction. (8 weeks.)
- Test streaming algorithm for 3D Morse-Smale complex construction. (8 weeks.)
- Implement parallel algorithm for 3D Morse-Smale complex construction. (8 weeks.)
- Test parallel algorithm for 3D Morse-Smale complex construction. (8 weeks.)
- Present results and obtain feedback from stakeholder. (1 day.)
- Choose best of streaming and parallel strategies and refine software and develop to deployment stage. (4 weeks.)

Feature tracking

- Develop strategy to handle increase in data-size due to increase in dimension.
  - Develop streaming algorithm for feature tracking. (3 weeks.)
  - Develop parallel algorithm for feature tracking. (3 weeks.)
- Implement streaming algorithm for feature tracking. (4 weeks.)
- Test streaming algorithm for feature tracking. (4 weeks.)
- Implement parallel algorithm for feature tracking. (4 weeks.)
- Test parallel algorithm for feature tracking. (4 weeks.)
- Present results and obtain feedback from stakeholder. (1 day.)
- Choose best of streaming and parallel strategies and refine software and develop to deployment stage.
  - Integrate 3D segmentation and tracking into visualization and interaction software already developed in stage A4. (4 weeks.)
  - Present software to stakeholder and get feedback. (1 day.)

Task A6. High quality visualizations and movies for publication and distribution.

- Iterate with scientists to produce high quality visualizations for publication as figures (2 weeks)
- Plan movies using low quality rendering (2 weeks)

- Render frames for movies using Povray or other high quality rendering package 50-200 hours per movie.

Risks and Opportunities

- OPPORTUNITY. In A4-Feature tracking we may be able to reuse in part or completely the code developed for gird combustion data. Likelihood: medium-high. Impact on project: high. Remediation: verify this possibility and adjust plan to speedup deployment.
- RISK. Transition from research prototype code to robust deployment code can be difficult. Likelihood: high. Impact on project: low. Remediation: We introduce several testing and feedback steps in each stage to ensure that the stakeholder is satisfied with the software. Moreover, all new development effort will use the best software engineering practices and focus on designing modular software. Affected tasks: A2 through A6.
- OPPORTUNITY. In A4 we may be able to reuse in part or completely the code developed for gird combustion data. Likelihood: medium. Impact on project: high. Remediation: verify this possibility and adjust plan to speedup deployment.
- RISK. Resource scalability: Large resource requirements to handle extremely large data-sizes. Likelihood: high. Impact on project: moderate. Remediation: We have access to several large high-performance computers in LLNL, LBL, and ORNL that can handle such data sizes. Affected tasks: A4.
- RISK. Software does not scale appropriate to increase in data-sizes. Likelihood: high. Impact on project: high. Remediation: We incorporate scalability as a design requirement in each of our tasks. In particular, we include a re-design of 3D Morse-Smale computation and feature software to handle large data-sizes. Moreover, we have experience in designing streaming algorithms that are very memory-efficient and can be adapted to the problems at hand. Affected tasks: A4.
- RISK Change in requirements from stakeholder. Likelihood: low. Impact on project: moderate. Remediation: We add feedback steps in each stage to identify any change in requirements as soon as possible. We keep our software designs sufficiently general and flexible to quickly add more functionality as and when required. Affected tasks: A2 through A5.
- RISK. Access to data requires accounts at LBL. Likelihood: high. Impact on project: moderate. Remediation: We have obtained accounts at NERSC computers and successfully downloaded sample datasets. Affected tasks: A2 through A5. **This item is no longer a risk factor and should be removed from the PMP.**
- RISK. The available computers do not have enough main memory to run the code for real data. Likelihood: low. Impact on project: high. Remediation: We have obtained accounts at NERSC computers and successfully downloaded sample datasets. Affected tasks: A2 through A5.

Accomplishments in this Reporting Period

- A1: Topology research presentation by Valerio Pascucci and meeting with John Bell and group at LBL.
- A2: In order to facilitate work on later stages, we used existing raw format scalar data and visualized it using in-house visualization code written using OpenGL.
- A3: There is some overlap between the feature tracking work in this stage and with similar work for the Sandia combustion analysis requirement. We have already started 2D feature segmentation and tracking for the Sandia requirement and will re-use the interaction framework developed for the LBL requirement.

- We had further discussions and refined the target of the work.
- We have started exchanging data and an exploratory evaluation of the existing software.
- We completed testing our 3D analysis code on the analysis of porous media. We are verifying how to use the same scheme for the combustion data.

Work Targets for the Next Six Months

Within three months:

- Finish A1.
- Work and test data for A2 and A3.
- Test evolving 3D Morse-Smale complex code.

Within six months:

- Work and test data for A4.
- Iterate at least two times with stakeholder to get feedback on quality* of segmentation into volume/tube/lines and preliminary tracking results.

Next 12-24 months.

Complete A3 and A4.

12.5 Fusion

12.5.1 Production Visual Data Analysis for Fusion

Team

- ORNL: Sean Ahern (team leader), Dave Pugmire.
- LLNL: Hank Childs.
- LBNL: Gunther Weber, Janet Jacobsen.

Stakeholder(s): John Cary, Tech-X Corporation.

Stakeholder Need(s):

Tech-X has identified a need for production quality visualization for fusion and accelerator simulation modeling delivered within VisIt. They have several different applications and have expressed the desire to have one fully-featured tool that meets their needs. They would like VisIt to be enhanced to meet this need.

This list of milestones and deliverables was generated by Sean Ahern during a visit to Tech-X in May of 2007. Gunther Weber, Janet Jacobsen, and Hank Childs participated in the visit as well. While ORNL has been able to address the first of their needs, we have yet to explore the delivery of the other milestones due to lack of personnel. With the addition of Dave Pugmire to the team, we now have the ability to start addressing needs.

Milestones/Deliverables/Objectives

Need A: Parameterized databases (datasets and index selection). VisIt has no ability for databases to take options from the user for tips in how to read the data. Tech-X would like such a capability.

- Task A1. We have implemented an initial version that allows for arbitrary groups of data to be requested from the user and passed to the database on the engine. (completed)

- Task A2. Much more could be added to VisIt for database options, so this deliverable needs to be assessed with Tech-X to see if it meets their needs.

Need B: Libsim in Vorpal

VisIt has the ability to view data from a running simulation, streamed to it from the simulation over a socket. Tech-X would like to integrate this into their Vorpal simulation code.

- Task B1. Needs assessment. Determine subset of Vorpal data that needs to be visualized.
- Task B2. Initial integration of libsim in Vorpal. Target: Summer 2008.

Need C: Better Streamlines

Description: The streamline plot in VisIt is somewhat rudimentary. A general need outlined by Tech-X personnel was improving this capability. Some specific requests follow as subneeds A-D:

Subneed C-A: Better interaction (tools)

Streamlines in VisIt are somewhat clumsy to interact with. Better tools for directly manipulating streamlines are required.

- Task C-A. Needs assessment. Identify exactly what interaction modes Tech-X would like.

Subneed C-B. Arbitrary streamline length.

VisIt's streamlines are limited in length. Arbitrary lengths (to limits of memory) are needed to do sophisticated analysis. Implementation of this likely requires good parallelization of the streamline algorithm.

- Task C-B1: Needs assessment. (completed)
- Task C-B2: Initial implementation of arbitrary length in scalar.
- Task C-B3: Parallel implementation of streamlines integration algorithm.

Subneed C-C: Streamline operator.

To allow the full range of VisIt analysis (including subsequent operators), streamlines must be output by an operator.

- Task C-C1: "Promote" VisIt's streamline plot into a streamline operator that can then be plotted by a pseudocolor plot (or other plots).

Subneed C-D: Integrators (symplectic)

VisIt's streamlines use 4th order Runge-Kutta integration. While good, these need to be augmented by additional integrators. Some suggestions were Adams-Bashforth and symplectic integrators that guarantee no integration across separatrices.

- Task C-D1: Needs assessment. We need to understand exactly what integrators are required and the calculus required.
- Task C-D2: Initial implementation in VisIt. This implementation may require modifications to the underlying VTK library.
- Task C-D3: Initial delivery. Target: Summer 2008.

Need D: Tabbed Windows

Many of Tech-X existing analysis tools use tabbed windows as an interaction modality. They would like the same economy of real estate to be implemented in VisIt's viewer windows. The primary task would be to implement tabbed windows in the VisIt viewer with a target start date of Spring 2008.

Need E: Better Keyframing (Camera Path Planning)

VisIt's keyframe editor is somewhat limited when it comes to planning the path of a camera. It is possible to choose certain views, then add those views to the set of keyframes. However, this interaction mode makes it difficult to plan camera flights through very confined spaces. Tech-X would like to have an on-screen editing mode where the camera path may be directly manipulated.

We have yet to consider the impacts or requirements for implementing this in VisIt. As such, the target date is at least a year out.

Custom GUIs

Tech-X's current data analysis tools provide custom interaction windows that are tailored to their application domain. They would like a similar capability in VisIt.

The beginning point for this task is a needs assessment: we need to determine exactly the range of GUI modifications required. Some of this may be met by VisIt's "new" Macro window.

Need G: Poincaré plots and analysis

Poincaré analysis allows for the understanding of magnetic field line winding and island identification. Once streamlines have been improved by subtask C-C above, such capability can be added to VisIt.

- Task G1. Needs analysis. Identify exactly what elements of Poincaré plots and analysis are required. Generate a prioritized list.
- Task G2. Initial implementation of highest priority items. Note: VACET team member Allen Sanderson (Utah) has already packaged much of SCIRun's Poincaré analysis into a stand-alone library that could conceivably be integrated in VisIt.

Need H: Islands extraction/plotting (from SCIRun for example)

Once the Poincaré plot generation from task G above is complete, island identification is needed. This is similar to what SCIRun does. We believe this will be a direct step from task G, using things like VisIt's threshold operator.

Need I: Simulation/Experiment comparison

Tech-X would like to do direct comparisons of simulation data to data gathered from experiment. While VisIt has the capability of doing this, specific features are required to meet all of Tech-X's needs.

Subneed I-A: MDSplus reader To compare MDSplus data, VisIt needs a native MDSplus reader.

- Task I-A1: Solicit a test dataset from Tech-X and implement a reader for this data.
- I-A2: Initial delivery of reader to Tech-X. Validate against other datasets.
- I-A3: Subsequent needs assessment.

Subneed I-B: Synthetic diagnostics To facilitate comparison, VisIt can generate synthetic diagnostics, simulating what would happen in a real experiment.

- Task I-B1: Needs analysis. Identify what synthetic diagnostics Tech-X requires. Prioritize list.
- Task I-B2: Initial implementation of highest-priority needs.

Need J: Better Launching on Laptops. Specifics of this need are not known at this time.

Need K: Resample average. VisIt's resample operator does not allow the ability to average the underlying data onto a new grid. Tech-X would like this ability. VisIt's ECF features allow this (and much more) capability. However, it's not easily accessible.

- Task K1: Make the ECF "average" ability accessible through the GUI.
- Task K2: Explore the concept of having the resample operator average the data when creating a new grid.

Need L: Particles advection VisIt has the ability to plot a particle field that is static in time. When applied to a time-varying field, this creates a sequence of plots that show moving particles. Similarly, VisIt has the ability to plot streamlines of a temporally static flow field. Tech-X would like the combination of these two. Specifically, they would like to have a plot that shows particles advecting through a field. This would require VisIt plots to be able to animate independently of the database time.

In addition, they would like this ability to be delivered for time-varying databases. Thus, particles would advect through a flow field that changes over simulation time.

- Task L1: Explore the addition of "plot animations" in VisIt.
- Task L2: Explore how to do a time-varying "plot animation" in VisIt.

Comment: Both of these are significant departures to how VisIt has operated in the past. Significant engineering may be required.

Need M: Switch between particles and grid quantities. Specifics of this need are not known at this time.

Need N: Plots with Temporal persistence. This is similar to task L above. Tech-X would like to have plots that have "temporal persistence," that is, plots that retain the data from timestep to simulation timestep. This would allow plots like streaklines, showing the data from a set of prior timesteps within one plot.

- Task N1: Needs assessment. Determine exactly what type of temporally persistent plots Tech-X would like to see. *Comment:* This task should really be done in concert with task L above, as the requirements for both are likely to be similar.

Need O: Multivariate volume rendering of species info. Specifics of this need are not known at this time.

Accomplishments

The primary accomplishments this period consist of a number of meetings and email exchanges between VACET and Tech-X to better understand the scope of activities required to realize the objective of "providing production-quality visual data analysis infrastructure for SciDAC fusion projects."

Also, we implemented an initial version of the database options in VisIt.

Work Targets for Next Six Months

Generally, we will be ramping up our interactions with Tech-X, doing the needs analysis for many of these tasks. Since the last large "project" communication we had was in May of 2007, we need to dive deep into the needs and prioritize the wants. That will be our initial tasks.

We hope to start making progress on streamline modifications in the next six months. We also hope to start initial work on major architecture changes for VisIt plots as outlined above.

Tasks beyond that will fall out from the needs analysis.

12.5.2 Interactive Visual Analysis of Large Fusion Particle-Based Datasets

VACET Team: Allen Sanderson, Utah (Team Leader); Lee Myers, Utah; Thiago Ize, Utah.

Customer/stakeholder(s): Stephane Ethier, PPPL. The SciDAC Center for Gyrokinetic Particle Simulations of Turbulent Transport in Burning Plasmas (W. Lee - PI).

Stakeholder need(s): The physicists are currently generating simulations that use millions to billions of particles, with each particle containing multiple scalar and vector data. They would like to have the ability to explore the nature of the particle orbits in an interactive manner. It is impractical to view billions let alone millions of particles at one time and glean any insight. As such, the physicists would like to have tools that allow them cull the particles with a user defined query or other statistical tools.

At the same time as particles are displayed physicists are in seeing the particles in context of data that is not associated with the particle but is part of the simulation. This data may be scalar (electric potentials) or vector data (magnetic field) and may have its own visualization requirements. For instance, the scalar data may be viewed using a variety of techniques from volume rendering to slicing.

Milestones/Deliverables

- Task A1. Creation of a GUI interface that allows for the efficient mapping of data values to glyph parameters.
- Task A2. Develop infrastructure with in SCIRun that allows for the processing of queries.
- Task A3. Utilize the Manta rendering environment for interactive exploration.
- Task A4. Develop infrastructure with in SCIRun that allows for interfacing with third party statistical tools.
- Task A5. Explore alternative tools, such as parallel coordinates for exploring the data.
- Task A6. In conjunction with the SciDAC SMD Center develop a data storage flow that allows for easy access by both physicists and visualization scientists that allows for post simulation analysis and visualization.

Each of the tasks are independent of one another.

Accomplishments

Our recent work has been to finalize a multivariate glyph that adequate represents the particle and its associated scalar and vector data. The glyph, a cone combined with a ring at the base has dubbed the "Sorcerer's Hat" (from Disney's Fantasia). This glyph can represent a vector and two scalar values. This work represents accomplishment of Task 1a (below).

The query tool implemented in SCIRun has been used to produced two "query based visualizations" of particle data. Through a small a sample size, each query was unique in that one was a simple query based directly on the particle data and only visualized those particles passing the

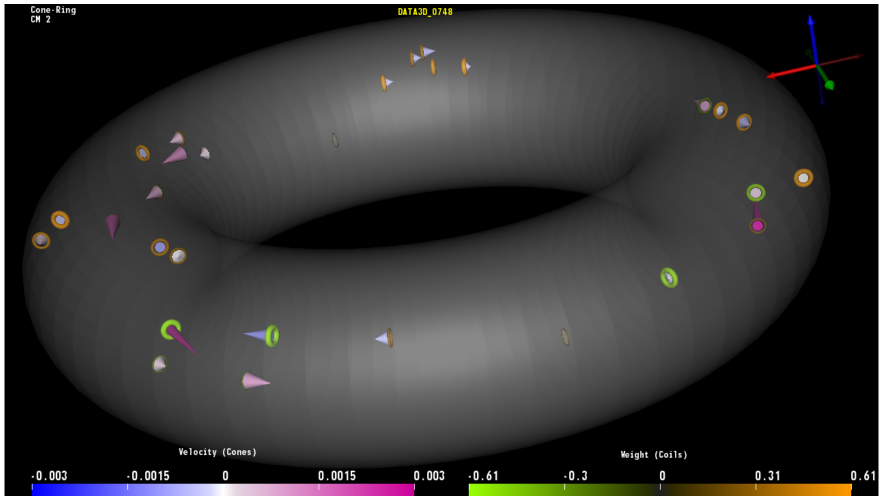

Figure 15: This image shows the new glyph-based technique applied to one timestep of particle-based data produced by GTC where two glyphs depict three data values associated with each particle: the velocity parallel to the magnetic field; the radius of the charged right about which the particle gyrates (the gyro-radius); and the statistical weight based upon departure from equilibrium. Here, we use a cone and a torus, each of which has two degrees of freedom, which are unified through the major radius value (and orientation). For the cone, the velocity magnitude maps to the length, and the the gyro-radii maps to the cone radius. For the torus, the gyro-radii maps to the major radius and the statistical weight maps to the minor radius.

query. The second was more complicated in that it required the calculation of intermediate data that was then used in the query. Further the results of the query where also used for visualizing not only the particles but other data associated with the simulation. This other (scalar) data was subsampled based on the particles being visualized. A movie of second query can be seen on the VACET website[29]. This work represents accomplishment of Task 1b (below) and is illustrated in Figure 16).

Work Targets for Next Six Months.

1. **Task 1a.** Enhancements to the SCIRun infrastructure for glyph based representations. This infrastructure is needed so that it will allow the mapping of multivariate data to each degree of freedom in the glyph. Currently much of this is hard coded and the number of variables that can be mapped is limited. (Completed Fall 2007)
2. **Task 1b.** Enhancements to the SCIRun infrastructure to create new visualization representations based on multiple times slices of data. When objects are in motion having visual cues help the observer envision the path. However this requires infrastructure in place that allows data at multiple time slices of data be processed at any given point. (Completed Fall 2007)
3. **Task 2.** Enhancements to the SCIRun infrastructure to allow queries. Query based visualization is a new area of work. The current infrastructure allows for single time slice pass-fail operations which is not acceptable with multiple time slices where coherence is needed.
4. **Task 3.** Expand the capabilities of our interactive renderer, Manta to more fully support glyphs and streamlining particles. By utilizing multiple CPUs for the rendering a more interactive setting can be achieved.

[29]http://www.vacet.org/gallery/#fusion

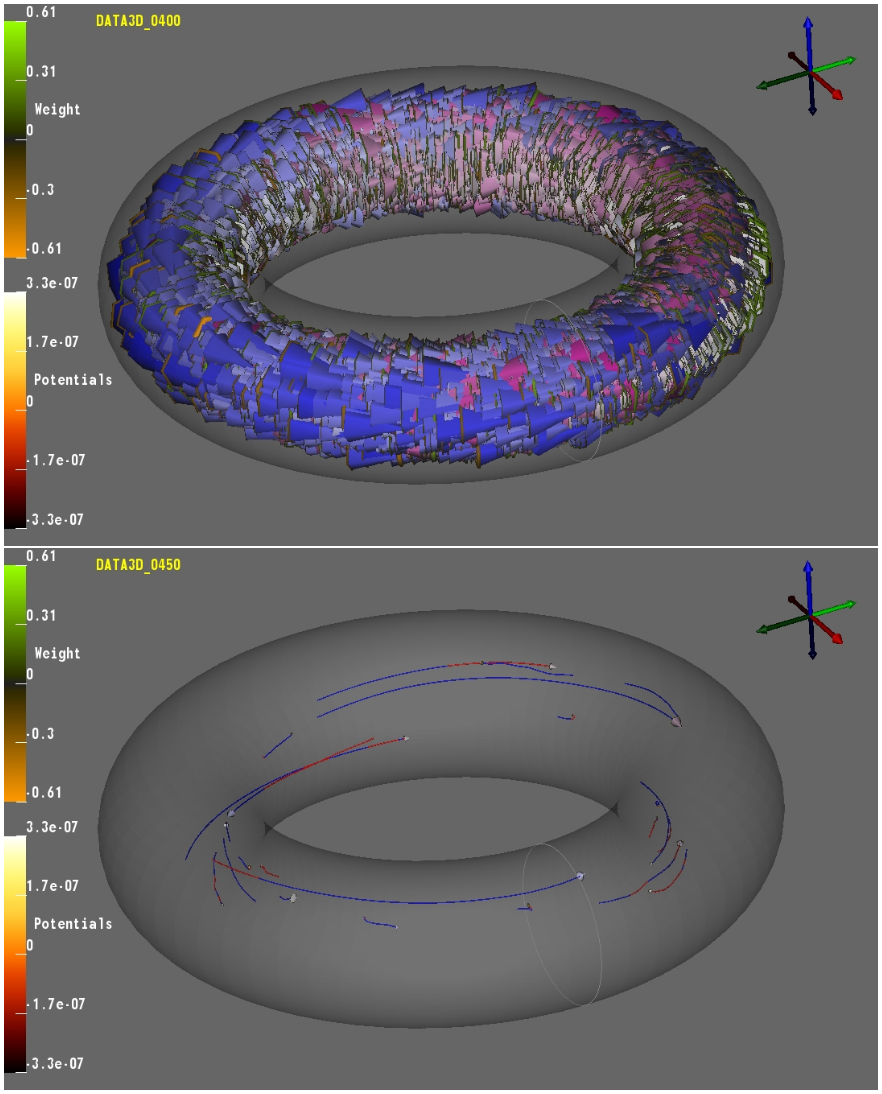

Figure 16: Query-driven visualization is a powerful metaphor for limiting visual data analysis processing only to "interesting data." Here, all particles that undergo trapping are shown in the top image, resulting in a high degree of visual clutter. Using our query-driven visualization machinery to restrict visual data analysis processing to an "interesting subset," we get a much more digestible image on the bottom. This bottom image shows the paths of about 50 trapped particles as they undergo trapping and untrapping, and use the glyph-based technique to convey additional information about each specific particle.

Risks

1. **Task 1a.** SCIRun GUI utilizes Tck/Tk which is cumbersome for complex GUI operations. Likelihood: low Impact: medium - multiple implementations may be required or limited capability Remediation: Utilize other tools for the GUI.
2. **Task 2.** The data flow infrastructure may not handle the query needs. Likelihood: medium Impact: requires a one off solution outside of the normal SCIRun paradigm. Remediation: Create an external tool.

3. **Task 3.** Additional glyphs (primitives) in Manta Likelihood:Low Impact: some glyphs may reduce performance Remediation: none.

Work Objectives Next Six Months

- **Task 4.** Develop infrastructure with in SCIRun that allows for interfacing with third party statistical tools. Statistical packages such as the R and GGobi packages provide powerful tools that can potentially allow clusters of particles to viewed. This could provide greater insight.
- **Task 6.** After the awarding of the SciDAC fusion projects we will re-examine this need in terms of the current funding.

Work Objectives Next 12-24 Months

Task 5. Explore alternative tools, such as parallel coordinates for exploring the data. We believe that viewing the data using an alternative coordinate system may provide insight not seen with traditional queries and statistical methods.

We fully anticipate that the needs will change as our interaction and deployment of tools to the physicists continues. For instance, the current early work has shown limitations in the way we access data stored as HDF5 files at the same time showing that the manner in which they store the data could be improved. This is necessary for the queries to have fast look up.

12.5.3 Poincaré Analysis and Visualization of Magnetohydrodynamics Simulation Data

VACET Team: Xavier Tricoche, Utah; Allen Sanderson, Utah (Team Co-leaders).

Customer/Stakeholder: Scott Kruger, Tech-X, working with the Center for Extended Magnetohydrodynamic Modeling (CEMM) SciDAC science application.

Customer/Stakeholder Need(s). Tools for automatic topological analysis of Poincaré plots in the study and assessment of Tokamak simulation data produced by NIMROD. The Poincaré plots are obtained by integrating field lines along a magnetic field.

Milestones/Deliverables

- **Task 1.** Deliver tool that extract significant topological information from fractal structures starting from discrete data. Additionally, the transient nature of the underlying phenomenon must be properly accounted for.
  - Write new software that adaptively integrate field lines in the magnetic field.
  - Apply edge detection to a scalar field derived from the field line information.
  - Extract X- and O-points using robust numerical search.
  - Reconstruct associated islands through solution of boundary value problem.
  - Measure relevant quantitative information associated with islands (number, volume, safety factor, ...).
  - Track the resulting structures over time using topological consistency to infer underlying continuous transformations.
  - Duration: 12 months. start date: October 1, 2006. completion: September 30, 2007.
- **Task 2.** Reliably handle the chaotic behavior exhibited by this type of data and convey it in the form of visual uncertainty in the resulting depiction.
  - Use topological structures to drive the integration of field lines in the direct vicinity of the island boundaries and X-points.

  - Apply to the resulting field lines coherence measure to quantify chaos.
  - Visualize this information in the form of a fuzzy color map superimposed on the topology.

  - Duration: 9 months. start date: October 1, 2006. completion – on hold.

- **Task 3.** Integrate tool directly in simulation code to guarantee full accuracy and optimal leverage of existing infrastructure.
  - Link Poincaré analysis code written in C++ to Nimrod Fortran code.
  - Duration: 2 months. start date: April 15, 2007. completion: on hold.

Task Dependencies

- Task 2 depends on Task 1 since chaos and uncertainty will be identified and quantified with respect to the know topological structure.
- Task 1 will benefit from Task 2 to assess the reliability of the computed measures.
- Task 3 may be concurrent with Task 1; both correspond to a non-visual data processing task.

Accomplishments

This project is currently "on hold" – the primary staff on the project, X. Tricoche, recently left the University of Utah for a position at Purdue University during the Summer of 2007.

Task 1.

- Item i. to iv. were already covered at the prototype level in the progress report of April 2007.

- The application of the prototype solution to the computation of magnetic islands in simulation data proved disappointing. The considered algorithm from the nonlinear dynamics literature had unforeseen shortcomings that made its application to NIMROD data unsatisfactory.
- An alternative approach was initiated that could not be completed before Xavier left the project. Since the extraction of the islands is pivotal for this project, a number of additional items remain on hold as well.

Task 2.

- Items i. to iii. were already covered in the previous progress report and have led to two accepted publications (the most recent one at IEEE Visualization 2007). Their application to the NIMROD data however was hampered by the difficulties faced in milestone 1 to properly characterize the geometric structures needed to guide the computation of structural stability that milestone 2 is concerned with.

Task 3. On hold due to departure of key staff.

Publications.

- C. Garth, G.-S. Li, X. Tricoche, C.D. Hansen. "Interactive Visualization of Coherent Structures in Transient Flows." In Topology-based Methods in Visualization, Mathematics + Visualization (TopoInVis), Grimma, Germany, March 2007.
- C. Garth, F. Gerhardt, X. Tricoche, H. Hagen. "Efficient Computation and Visualization of Coherent Structures in Fluid Flow Applications." IEEE Transactions on Visualization and Computer Graphics 13(5) (Proceedings of IEEE Visualization 2007), October 2007.

Work Targets for Next Six Months

The work is on hold due to the departure of key staff.

Work Targets Next 12-24 Months

The work is on hold due to the departure of key staff.

12.6 Mathematics

12.6.1 Production-Quality AMR Visual Analysis Infrastructure

VACET Team Members

- LBNL: G. Weber (Team Leader), W. Bethel
- LLNL: H. Childs, B. Whitlock, K. Bonnell
- ORNL: J. Meredith
- Related VACET team members from efforts for volume rendering, line integral convolution, streamlines, and embedded boundaries.

Customer/Stakeholder(s)

- Phil Colella, LBNL and the SciDAC Applied Partial Differential Equations Center (APDEC).
- John Bell, Center for Computational Sciences and Engineering (CCSE), LBNL; S. Woosley's Computational Astrophysics Consortium SciDAC Science Application.
- Ravi Samtaney, PPPL; Center for Extended Magnetohydrodynamics via APDEC

Stakeholder Need(s)

Both primary stakeholder teams – APDEC and CCSE – have a need for production-quality AMR visual data analysis infrastructure. Both teams have an in-house tool they have been using and funding over the years. Both teams have expressed the urgent need to get out of the business of developing and maintaining such software infrastructure. Additionally, both teams have expressed the need for new capabilities that lie outside the scope of what is possible with their existing technology. The activities to support these teams can be described with four high level thrusts:

- A. Invest in critical, missing visualization algorithms
- B. Planned interface changes to ease the transition of their user community
- C. Planned functionality improvements specific to APDEC
- D. Maintenance of software infrastructure, bug fixes, etc.

A more detailed list of specific needs is documented on the VACET wiki. That list contains around forty different features requested by customers from APDEC and the Astrophysics SAP project. Because the tasks are so numerous and since many will take a significant amount of time to accomplish, our approach is to prioritize these features to optimize the intersection between (1) the stakeholders assessment of the features importance and (2) the degree of difficulty in implementing the feature. With this prioritization, we order our work to achieve impact quickly while continuing work on more difficult and challenging research and engineering problems. The tasks of each of the four categories are itemized below along with labor efforts, stakeholder priority, and status (in some cases where appropriate).

Milestones/Deliverables/Objectives

A. Invest in critical, missing visualization algorithms

- **Task A.1.** (High priority) Fix slow performance on rectilinear AMR grids. Estimated time: 4 weeks. Status: completed as of April 2007.
- **Task A.2.** (High priority) ChomboVis-style visualization spreadsheets capability. A detailed description is posted on the VACET wiki. Estimated time: 6 weeks. Status: completed as of April 2007.
- **Task A.3.** (High priority) AMRVis-style visualization spreadsheets capability. A detailed description is posted on the VACET wiki. Estimated time: 6 weeks.
- **Task A.4.** (High priority) Invoke VisIt from the debugger. The primary work here is to write a VisIt component that connects to VisIt's main state (in the "viewer") and is able to communicate with GDB. The current implementation (AMRVis/ChomboVis) requires starting up the tool for each new patch. We would like to avoid that here, especially because there have been complaints about startup time for the existing solutions and VisIt not do any better in terms of startup time. Estimated time: five weeks. This work is described in more detail on the VACET wiki as well as in this PMP in Section 12.6.2. Status: completed as of October 2007.
- **Task A.5.** (High/medium priority) Familiar, customizable interfaces. The approach is to use a ".visitrc" file so that the UI can be customized to appear as either AMRvis or ChomboVis. Estimate: six weeks. Status: completed as of October 2007.
- **Task A.6.** (High priority) Embedded boundaries. APDEC has said they will not switch entirely to VisIt until there is embedded boundary support in VisIt. ChomboVis has an existing solution. Work here is underway with VACET team members from UCD to generate a rigorous mathematical solution that would eventually find its way into VisIt. That work is described in more detail on the VACET wiki as well as in this PMP (Section 12.6.3).
- **Task A.7.** VisIt's ghost data generation works well for some AMR cases, but not others. This algorithm needs to be improved. Estimate: 6 weeks.
- **Task A.8.** (High priority) Streamlines. APDEC would like a better implementation of streamlines in VisIt. Specifically a good parallel implementation that crosses patch boundaries and allows for different integration schemes. That work is described in more detail on the VACET wiki as well as in this PMP (Section 12.6.4).
- **Task A.9.** (High priority) Enhancement: It takes too much work to get an outline of the AMR patches. This isn't as trivial as it sounds, because the data travels through a data flow network, and you have to be able to still draw outlines after contours, slices, etc. Estimate: two weeks. Status: completed as of October 2007.
- **Task A.10.** (Medium priority) Better volume rendering. This milestone is the subject of a significant amount of VACET work as described on the VACET wiki as well as in this PMP (Section 12.7.1).
- **Task A.11.** (Medium priority) 3D LIC/Vector field visualization. This feature requires a long lead time due to the need to develop new technology combined with significant engineering challenges. This effort is described in more detail on the VACET wiki as well as in this PMP (Section 12.10.1).
- **Task A.12.** (Medium priority) Face- and edge-centered data. (cost needs to be assessed; 6 months+ for complete solution).
- **Task A.13.** (Low priority) User-supplied callbacks. For both expressions (i.e., derived quantities) and queries (i.e., data analysis routines). Bell wants to be able to provide his own callbacks that perform expensive physics computations. (cost needs to be assessed; 3 month estimate).

B. Planned interface changes to ease the transition of their user community

- **Task B.1.** (High priority) Connect slice and clip operators. Customers would like to be able to modify one of these operator and have the other automatically update. This was done by linking them through VisIt's plane tool. Estimated time: two weeks. Status: completed as of October 2007.
- **Task B.2.** (High priority) Mouse scroll wheel. Customers would like to be able to use the mouse scroll wheel to do zoom operators. Estimated time: two days. Status: completed as of October 2007.
- **Task B.3.** (High priority) Linking picks and visual spreadsheets. Customers would like for picks to automatically bring up visual spreadsheets, focusing on the cell of the pick. Estimated time: two weeks. Status: completed as of October 2007.
- **Task B.4.** (High priority) Highlight picks. Customers would like for picks to highlight the surrounding cell and its patch. Estimated time: two weeks.
- **Task B.5.** (High priority) Slider for animating through slices. This should be easy, but the slider needs to make sure that each advancement corresponds to one logical index, and that index information is not accessible in VisIts GUI. So the real work here is to add that meta-data information. Estimated time: two weeks.
- **Task B.6.** (Medium priority) Grouping plots. Customers would like a mechanism in the user interface to group plots into a set that can be jointly manipulated with respect to operators, subset selection, etc. This represents a significant user interface change. Estimated time: one month.
- **Task B.7.** (Low priority) Automatic re-picking. Customers would like to be able to indicate a variable of interest and have picks automatically update with that variable information. Estimated time: three days. Status: completed as of October 2007.
- **Task B.8.** (Priority: TBD) Changes to view interactors. Customers would like to be able to customize the way the view is manipulated (e.g. left mouse button performs a pan, etc.). Estimated time: two weeks.
- **Task B.9.** (High priority) Selecting logical coordinates. This feature maps well to VisIt's IndexSelect operator, but that operator would have to be made AMR-aware (to state what level the logical indices are referring to). The slider needs to make sure that each advancement corresponds to one logical index, and that index information is not accessible in VisIts GUI. The bulk of the work here is to add the needed meta-data information. Estimated time: two weeks. Status: completed as of October 2007.

C. Planned functionality improvements specific to APDEC

- **Task C.1.** (High priority) Link Chombo's embedded boundary information into VisIt's material interface reconstruction algorithms as a short term measure. Estimated time: two weeks. Status: completed as of October 2007.
- **Task C.2.** (High priority) Use the ghost data outputted by Chombo, rather than generating it inside of VisIt. Estimated time: two weeks.
- textbfTask C.3. (High priority) Support Chombo's new file format for embedded boundaries. Estimated time: two weeks.
- **Task C.4.** (High priority) Mapped grids (Curvilinear AMR grids). While VisIt already supports curvilinear grids, including curvilinear AMR grids, the Chombo file format for curvilinear grids is presently in flux. Estimate: three days.
- **Task C.5.** (High priority) Have VisIt only bring up refinement level zero by default to less processing time. This requires new VisIt infrastructure. Estimated time: one week.

- **Task C.6.** (High priority) Have VisIt only bring up specific materials by default. This requires new VisIt infrastructure. Estimated time: one week.
- **Task C.7.** (High priority) Allow Chombo to define derived quantities that automatically show up in the user interface. Estimated time: two days.
- **Task C.8.** (Medium priority) Support double precision data. Double precision data is often not needed for common visualization operations, so more information is needed to determine exactly what areas this is necessary for. Estimated time: one week (partial solution).
- **Task C.9.** (High priority) Ensure "new" VisIt runs on platforms important to Stakeholders. This task is expected to be without substantial complications since VisIt has been ported to work on essentially all modern parallel platforms. Estimated time: two weeks. Status: completed, yet also ongoing.

D. Maintenance of software infrastructure, bug fixes, etc.

Obviously, this is a highly dynamic list. To date, the following issues have been pointed out by stakeholders (primarily Brian van Straalen and Terry Ligocki).

- **Task D.1.** Missing boundaries. In 3D Mesh, Pseudocolor etc. plots, meshes can have missing boundaries. This seems to occur if a boundary is only partially shared with a patch from the same level. In that case, the entire boundary surface is discarded instead of only the part abutting to the same level patch.
- **Task D.2.** Inconsistent material selection. BVS and TL demonstrated a bug with materials being selected on some patches but not others that says "contact a VisIt developer". This is related to VisIt creating ghost data. If we were using Chombo ghost data instead, this bug would not come up.
- **Task D.3.** Line stool streamline problem. Status: completed as of October 2007.
- **Task D.4.** Scatter plot labels problem: axis labels not correct for the scatter plot and for the elevation operator.
- **Task D.5.** Visitrc problem: Visit sometimes only brings up an error dialog instead of loading macros if a visitrc file exists.
- **Task D.6.** Network problem: Brian's machine stops working when disconnected from the network.

There are six defects, D.1-D.6, that have been reported by the APDEC team. One, D.3, was resolved as of October, 2007. The remaining five have estimates of two to three days each to fix.

Accomplishments

- Initial conversations with the APDEC team led to a gap analysis of missing features in VisIt. These features had to be added for their team to transition away from ChomboVis. And, over the last reporting period, **VACET team members implemented all of the features from this initial gap analysis**. The points of contact for the Chombo team (Brian van Straalen and Terry Ligocki) **have transitioned to using VisIt** for the majority of the time and are poised to roll VisIt out to the entire team. In addition, many external collaborators have already begun using VisIt instead of ChomboVis.
- Regular (almost daily) meetings occur between the Chombo points of contact and VACET personnel (usually Gunther Weber). From these meetings, a new list of requirements were formed, which are represented in the revised PMP above.

- Specific tasks completed were: ability to connect to VisIt from a gdb debugging session of Chombo (A4), ability to customize the VisIt interface, done through new “macro” functionality (A5), increased performance, specifically for creating renderings of patch boundaries (A9), customization of the VisIt user interface to ease migration (B1, B2, B3), ability to select what data is read in through logical coordinates (B9), basic support for embedded boundaries (C1), and various bug fixes (D3).
- **Publications.**
  - June 2007, “Visualization Tools for Adaptive Mesh Refinement Data” [56]. This survey paper, authored jointly by representatives from VACET and APDEC, sketches the landscape of current AMR visualization capabilities in various tools, with a focus on the special visualization challenges posed by AMR data.
- **Presentations.**
  - June 2007, “Visualization Tools for Adaptive Mesh Refinement Data.” G. Weber, LBNL, at the 2nd International Conference on Numerical Modeling of Space Plasma Flows ASTRONUM-2007 (Paris, France, June 11-15, 2007).
  - June 2007, “Visualization Tools for Adaptive Mesh Refinement Data.” G. Weber, LBNL, the 4th High-End Visualization Workshop (University Center Obergurgl, Tyrol, Austria, June 17-22, 2007).

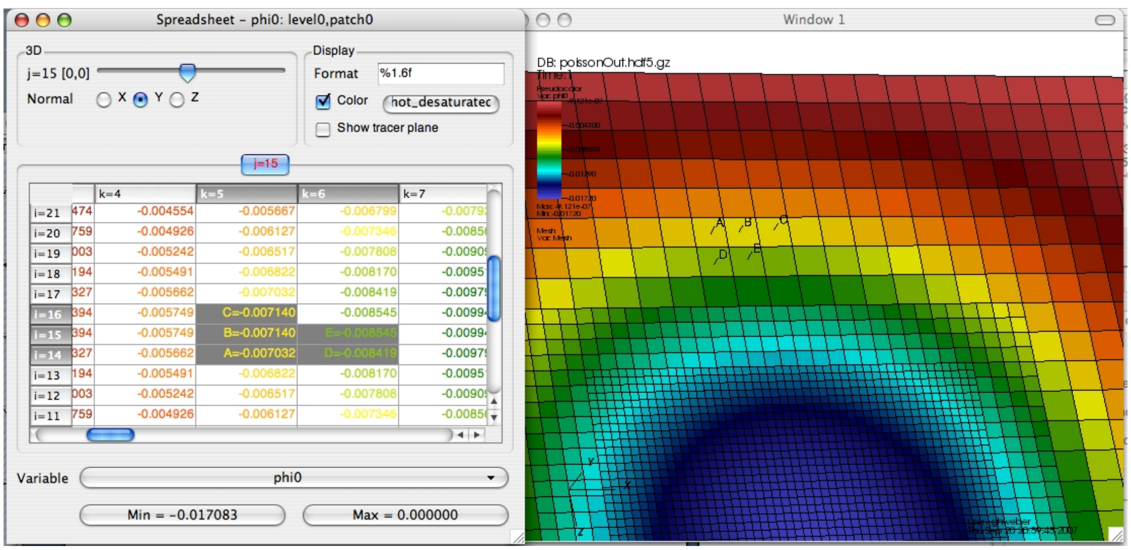

Figure 17: This image is a screenshot of the linking between picking and visual spreadsheets, task B.3. This capability was requested by our stakeholders, developed and released in production form in October 2007.

Work Targets for the Next Six Months

- B4, B5, B7, B8, C2, C3, C4, C5, C6, C7, D1, D2, D4, D5, and D6

Risks

- Many of the planned changes do not involve high risk. They are interface changes to ease the transition of the Chombo team or straightforward bug fixes.

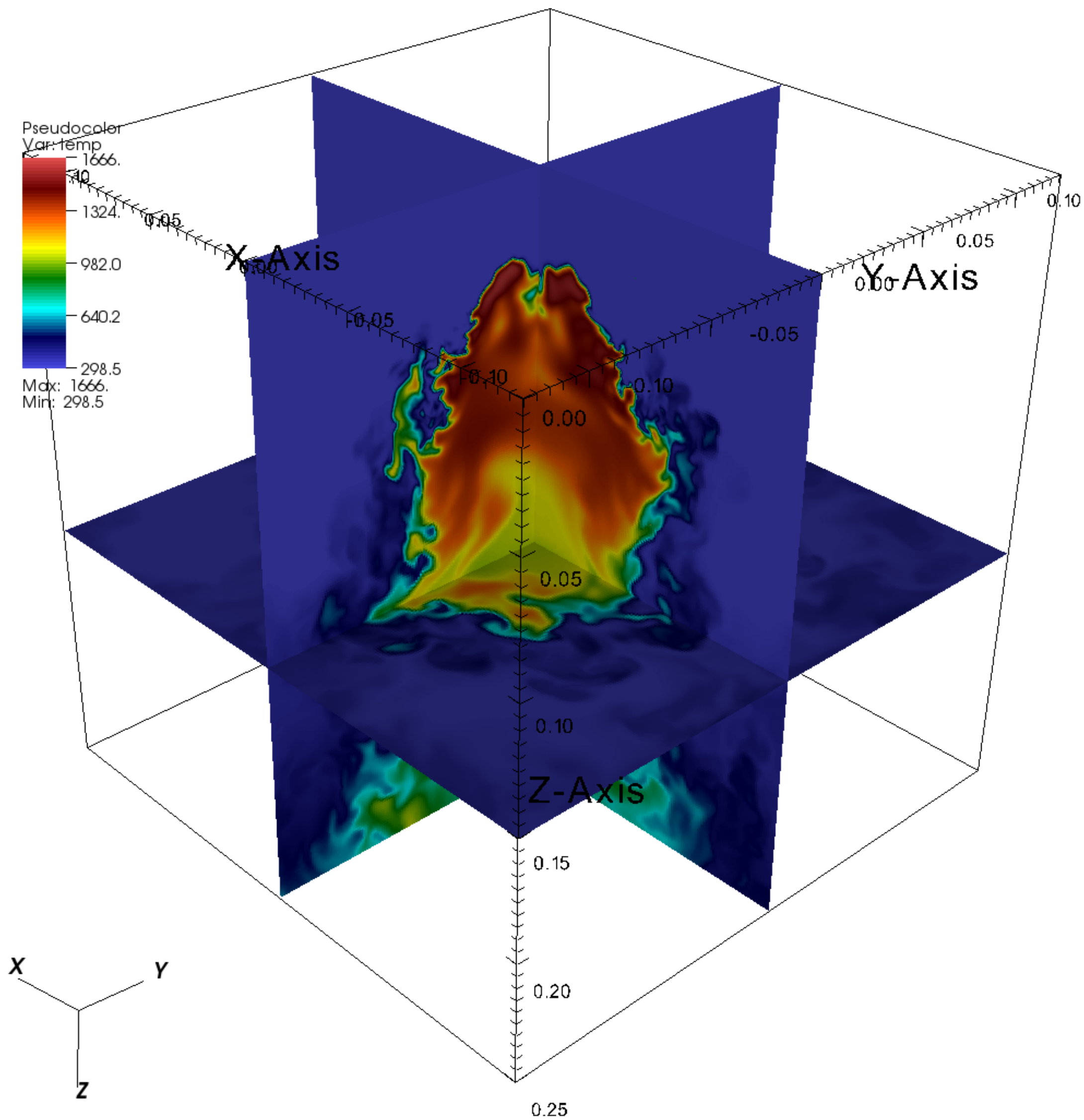

Figure 18: Orthogonal slices of a 3D AMR dataset. Sample data is a simulation of a H_2 flame, data courtesy of J. Bell and M. Day, LBNL/CCSE.

- The work targets for the next six months are likely to be dynamic, as the Chombo team's transition to VisIt will likely lead to many new requirements.
- The tasks for APDEC only total approximately 16 weeks. The personnel assigned to work on APDEC work are also pursuing A6, A8, A10, & A11, each of which is a high priority item with a large cost. There is a high likelihood of insufficient resources.

Work Plans Next 12-24 Months

- A.3, A.7, A.13, and B.6
- There will likely be many new requirements when VisIt is used widely throughout the SciDAC's AMR community.

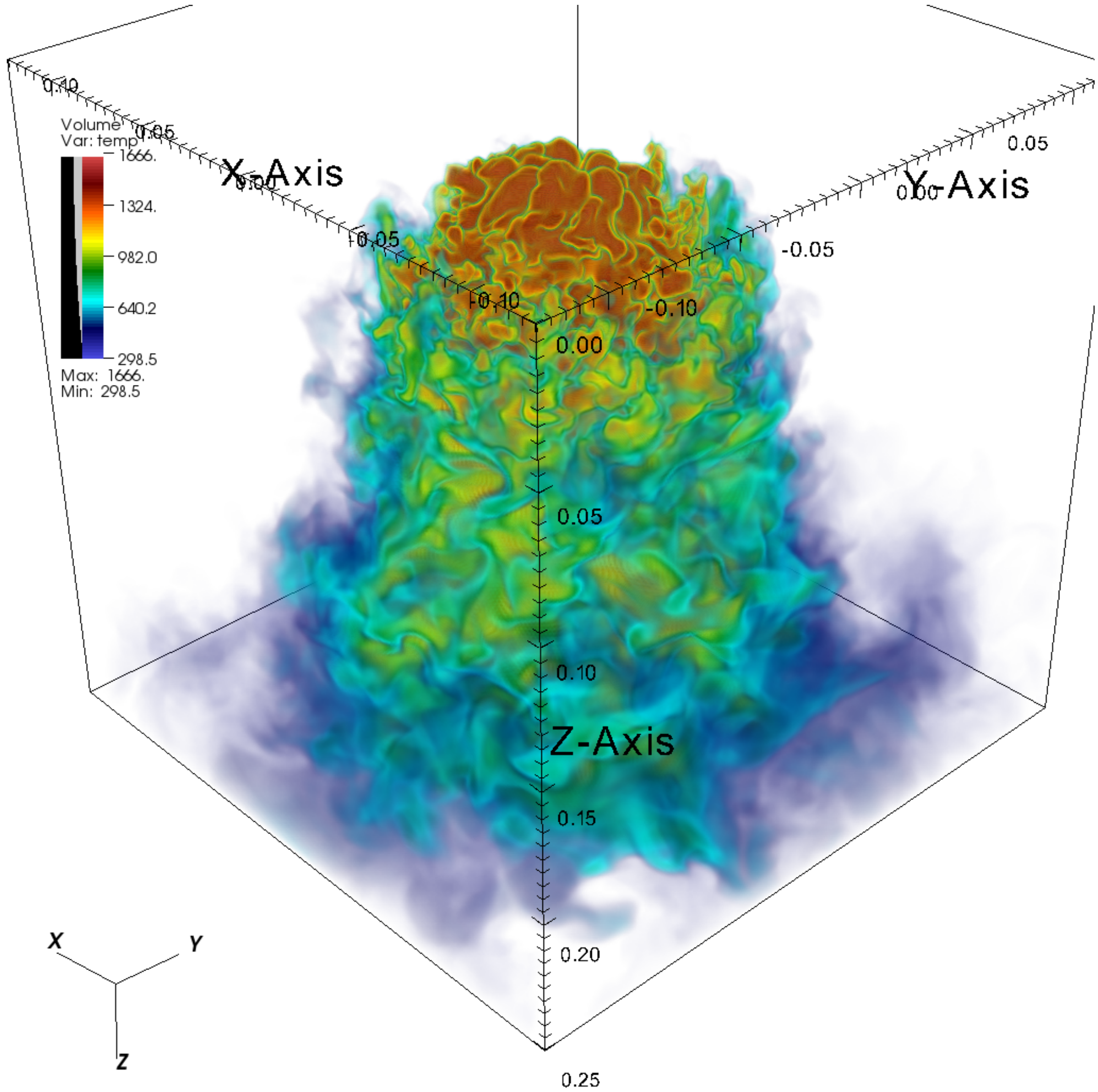

Figure 19: Volume rendering of a 3D AMR dataset. Sample data is a simulation of a H_2 flame, data courtesy of J. Bell and M. Day, LBNL/CCSE.

12.6.2 Debug With VisIt

Note: this section drills more deeply into one of the milestones listed in Section 12.6.1, namely being able to invoke a visual analysis application, VisIt, from a debugger while running a simulation code. Please refer back to Section 12.6.1 for details about the Stakeholder, VACET team members, and so forth.

Currently, Chombo can use the ChomboVis and ChomboBrowser visualization/analysis applications from inside a debugger like GDB. Literally, one sets a breakpoint at a crucial location, and then executes the GDB function "call" to invoke an internal routine with a pointer to some Chombo data structure to be visualized. An example might go something like this (all text simplified from its actual output):

```
> break main:3000
> run
breakpoint at main:3000
> l
3000   Grid *my_grid = FirstIteration();
> call viewLevel(my_grid) \\(saving temporary file; invoking process chombovis)
```

One invokes ChomboVis to view levels/blocks and ChomboBrowser to explore them numerically. The downside to this path are several: new ChomboVis processes take time to launch; the reliance on ChomboVis adds maintenance costs for those developers; ChomboBrowser and ChomboVis can't interact, nor can multiple ChomboVis invocations.

The planned (and now implemented) path is to replace this functionality with VisIt, with the following changes:

- VisIt is used for both visualization and browsing through the Pseudocolor and Spreadsheet plots, respectively.
- A single visit process handles all requests for visualizing or browsing.
- The first invocation is the only one which launches VisIt.
- Remaining invocations control VisIt through a hidden python interface (using UNIX pipes) .

Milestones/Deliverables

- **Task 1.** Supplant existing ChomboVis debugging infrastructure with VisIt.
  1. Launch VisIt using command from debugger.
  2. Support creation of spreadsheet plots in VisIt through the debugger using data from executable.
  3. Support creation of pseudocolor plots in VisIt through the debugger using data from executable.
- **Task 2.** Improve primary usability obstacles.
  1. Improve startup time (e.g. leave VisIt running and amortize startup across all debugging tasks).
- **Task 3.** Further improve usability.
  1. Support new plot types.
  2. Improve integration between running VisIt and running debugger (e.g. new GUI interface).

Dependencies

- Task 1 must occur before Tasks 2 and 3.
- Task 2-1 must occur after Task 1-1, but may be tied to implementation of Tasks 1-2 and 1-3.
- All Task 3 milestones are speculative, and as potentially unnecessary, should occur after all Task 1 and 2 objectives have been met.

Accomplishments

All milestones for Tasks A1-A3 and Task B1 completed and put into production deployment.

Work Targets for Next Six Months

Speculatively, Task C may become a requirement depending on end-user feedback, necessitating milestones C1,C2.

- Expected start date: January 2008.
- Expected duration: 1 month.
- Risks: none anticipated.

Next 12-24 Months

None.

12.6.3 Embedded Boundaries

Team

VACET: John Anderson, UCD; Hank Childs, LLNL; E. Wes Bethel, LBNL; Kenneth I. Joy, UCD (Team Leader)

Non-VACET: Mark A. Duchaineau, LLNL.

Stakeholder(s): Phil Colella/APDEC, LBNL, producers of AMR-based codes that utilize volume fractions (along with other quantities) in volume-of-fluid applications; Mark A. Duchaineau, LLNL, who studies neutron transport visualization problems ; VisIt Users where material boundary calculations are a feature of VisIt and heavily used.

Stakeholder Need(s):

In many applications it is necessary to reconstruct or track the boundary surfaces, or "interfaces," between multiple materials that commonly result from multi-fluid Lagrangian-Eulerian hydrodynamics calculations. This problem, where the generated data sets have the characteristic that each cell contains a "fraction" representing the percentage of each material contained in the cell, now arises in a variety of applications, and is frequently called the **embedded boundary problem**. The challenge is to utilize the material fractions in each cell to reconstruct the boundaries between materials. VACET researchers have been developing a number of possible solutions to this problem.

Having a production-quality solution to this problem will allow one of our primary stakeholders and his team (Colella/APDEC) to completely adopt VisIt for their visual data analysis infrastructure. The benefit to their team is a substantial cost savings – they will no longer have to create, update and maintain their own code base for visual data analysis.

Several versions of these methods must be produced:

- A generic solver whose input consists of volume fractions only. Up to eighty materials per simulation may be present.
- A two-material solver, where input consists of volume fractions and perhaps other quantities (area fractions, centroids, etc.) This solver needs to be accurate to within a pre-specified error tolerance.
- A "quick" solver for both of the above that minimizes computational time.
- A solver that utilizes a pre-specified geometry budget.
- A solver that runs in parallel and can be integrated into the VisIt framework.

Milestones/Deliverables/Objectives

In solving this problem, we must overcome several difficulties (1) the original volume fractions must be preserved, (2) the solution must be scalable, (3) the solution must be computationally "quick", and (4) adaptive techniques must be developed that can be tailored to the user's requirements. We will implement this into VisIt to get the maximum impact with the scientists.

- **Task A1.** Survey the national laboratory community to determine the scope of this problem, developing further tasks and milestones as to the needs of our stakeholders. (Complete).
- **Task A2.** Develop a framework that identifies the possible topologies of the embedded boundary solution.
- **Task A3.** Generate the two-material solver, utilizing only volume fractions. Determine and quantify the deficiencies of the algorithm. Enhance the basic solver with additional information (centroids, etc.) to produce more accurate results.
- **Task A4.** Generate the generic solver, whose input consists of volume fractions only. Generate the parallelizable solver.
- **Task A5.** Modify these solutions to produce adaptable versions of the algorithm that can be utilized to produce "near accurate" solutions under time constraints, and geometry budgets.
- **Task A6.** Integrate these solutions into VisIt.

Dependencies

- A2-A6 requires the completion of A1
- A3 requires the completion of A1 and A2
- A4 requires the completion of A2 and partially A3
- A5 requires the completion of A4
- A6 requires the completion of A5

Subtasks

Task A1. Work Team: C. Garth, K. Joy. **Completed.**
Task A2. Work Team: J. Anderson, C. Garth, K. Joy

- The general problem that inputs only volume fractions is significantly under-constrained, we need to determine all possible topologies that can be produced by these algorithms, quantifying them to determine further constraints on our solver.
- Determine parameters that can be used to influence the topological solution to the problem.

Task A3. Work Team: R. Martinez, J. Anderson, C. Garth, K. Joy

- Generate a two-material solver, using volume fractions as input. Utilize gradient methods to insure all boundaries are discovered and displayed. Iterative methods will be used to generate an adaptable solution.

Task A4. Work Team: R. Martinez, J. Anderson, C. Garth, K. Joy

- Generate the general n-material solver. Again, iterative methods will be employed.
- Develop the parallelizable solver.

Task A5. Work Team: R. Martinez, J. Anderson, C. Garth, K. Joy

- Develop the adaptable methods that develop solutions within given geometry budgets.
- Develop the adaptable methods that develop solutions within time constraints.
- Develop the adaptable methods that develop solutions within accuracy constraints.

Task A6. Work Team: R. Martinez, C. Garth, H. Childs, K. Joy

- Integrate these methods into VisIt.

Accomplishments

Task A1 Completed – We have surveyed our national laboratory stakeholders and developed a new task line that will enable the development of solutions that will maximally impact their work.

- LBNL/APDEC requires a two-material solution to the problem. Input could be a variety of volume fractions, centroids, area fractions and normal vectors. They could possibly require implicit methods to be used.
- LLNL requires a multi-material solution to the problem (80 materials maximum), where only volume fractions are input.
- LANL requires a multi-material solution, where centroids, area fractions and volume fractions are input.
- ORNL requires the same solutions as LLNL.

Task A2 In Progress We have developed new methods to determine the possible topologies for embedded boundaries. These methods are currently underdevelopment.

We are developing a dual-contouring approach that enables adjustment of initial boundaries to conserve material fractions. We have also developed a mathematically-defined guidance surface that enables triangles of varying sizes to be generated by the algorithm and allows interface reconstruction with specified error bounds on the volume fractions.

Project page: http://www.idav.ucdavis.edu/ joy/VACET/EmbeddedBoundary.pdf.

Work Targets for the next Six Months

Task A2. Work in progress, duration six months

Task A3. Work in progress, duration six months

12.6.4 AMR Streamlines Visualization

Team

- LBNL: G.H. Weber (Team Leader), B. Van Straalen (participating APDEC member), and T.J. Ligocki (participating APDEC member).
- UC Davis: L. Feng, and C. Garth.
- ORNL: S. Ahern and D. Pugmire.
- LLNL: H. Childs.

Stakeholder(s): The primary stakeholder is P. Collela, LBNL/APDEC, and J. Bell, LBNL/Computational Astrophysics Consortium, both of whom develop and apply AMR-based codes to a diverse array of scientific research problems.

The work from this project will feed into our Production Quality AMR visualization software infrastructure (see Section 12.6.1 and will likely result in new visualization publications since this work represents new, fundamental algorithmic design and implementation.

Stakeholder Need(s):

Currently, VisIt's streamlines have some deficiencies when applied to AMR data:

- They do not cross patch boundaries.
- They do not have the desired level of continuity at boundaries.
- There is little information as to streamline quality in general and in particular when streamlines cross patches.

Furthermore, VisIt's streamlines need to be improved to allow for:

- Parallel streamline computation.
- Treatment of multiblock/distributed datasets.
- Pathlines (time-dependent datasets).
- Support for multiple/custom Initial Value Problem (IVP) solver types with advanced features (e.g. continuous output).
- Use of streamlines as building-blocks for advanced visualization algorithms.

Milestones/Deliverables/Objectives

- **Task A1.** Reorganize VisIt streamline integration code to allow fine-grained access to streamline functionality.
- **Task A2.** Create parallel/multi-block streamline computation algorithms that are independent of the underlying integration scheme.
- **Task A3.** Enable ChomboVis equivalent functionality in VisIt:
  - Implement streamline support at the same level as ChomboVis in serial mode. ChomboVis currently uses the approach: Place sample. Find finest-level patch containing that position (using original patch outlines).
  - Use dual-grid and component-wise bi-/trilinear interpolation to obtain vector at sample position (utilizing ghost cell information to ensure that dual mesh encompasses all fine grid positions at the boundary).
- **Task A4.** Develop metrics for the evaluation of streamline integration quality in the presence of dataset irregularities (especially AMR resolution boundaries) and use them to compare multiple interpolation/IVP schemes.
- **Task A5.** Deploy in VisIt.

Dependencies

- A2 requires partial completion of A1.
- A3 requires total completion of the set of things in A2, etc.
- A4 can be started in parallel with A3, but requires A3 for completion.
- A5 requires completion of all other milestones.

Subtasks

Task A1. Work team: L.Feng, G. Weber, C. Garth.

- Introduce new IVP solver class hierarchy into VisIt with the aim of implementing individual IVP schemes as subclasses.
- Incorporate several integration schemes such as RKF45 and DOPRI56 (and possibly tie in IVP libraries such as CVODE).
- Introduce streamline (convenience) class more specifically aimed at streamline integration requirements (i.e. construction and storage of a global IVP solution) and a matching datatype to pass streamline sets between VisIt filters and plugins.

Task A2. Work team: S. Ahern, D. Pugmire.

- Parallelize IVP solver and streamlines. Communication issues.
- Adapt streamline integration to multiblock datasets.

Task A3. Work team: L. Feng, G.H. Weber, C. Garth.

- Use dual grid interpolation and ensure presence of proper ghost zone information.
- Get ChomboVis-equivalent serial version to work that transitions between patches of AMR hierarchy as necessary.

Task A4. Work team: B.van Straalen, T. Ligocki, G.H. Weber, C. Garth, L. Feng.

- One possible metric: Seed points along shape (circle, line, etc.). Calculate streamline of certain length for each seed point. Evaluate, how much shape of resulting end points deviates from original shape. Look at other metric. Look how decrease in step size affects error. (Requires ANAG/APDEC input).
- Integrate these metrics in VisIt. For doing that, we need infrastructure to, e.g., query the positions of the last points of a couple of streamlines (available from A1).
- Identify interpolation schemes.
- Identify possible IVP scheme candidates.
- Implement and evaluate using metrics.

Task A5. Work team: all except T. Ligocki and B. van Straalen.

- Harden code-base for final "chosen" candidate(s).
- Gradually replace currently used VTK classes with new streamline code.
- Add regression tests.

Duration, Start Date, Completion Date

- **Task A1.** Start now. Duration: 2-3 months.
- **Task A2.** Insufficient information at this time to estimate.
- **Task A3.** Start: 12/2007 or 01/2008. Duration: 2-3 months.
- **Task A4.** Start: 12/2007 or 01/2008. Duration: 2-3 months.
- **Task A5.** Start: 05/2008. Duration: 1 month.

Accomplishments

The primary accomplishments for this new task are: (1) requirements discussions with the stakeholders to determine needs; (2) initial analysis of requirements and preliminary design; (3) documentation of projected work plans for this project.

Work Targets for Next Six Months

Task A1. Work in progress. Duration: 2-3 months.

Task A2. Insufficient information at this time.

Risk: Approach developed in A1 not parallelizable. Likelihood: low (if there is communication between teams). Impact: high. Remediation: Develop this scheme for existing streamline functionality independently from AMR and use lessons learned for A3/A4.

Task A3. Start: 12/2007 or 01/2008. Duration: 2-3 months.

Risk: ChomboVis approach does not map well to VisIt. Likelihood: low. Impact: high. Remediation: Examine equivalent approaches enabled by A1.

Risk: Necessary queries difficult to implement in VisIt. Likelihood: low (due to work in A1). Impact: medium. Remediation: Write out streamlines and use external modules. If there are other

applications for certain queries apart from streamline quality measures, evaluate trade-off in time necessary to implement them anyway and utility gained.

Task A4. Start: 12/2007 or 01/2008. Duration: 2-3 months.

Risk: Developed schemes do not fit in VisIt framework. Likelihood: low-high (depending on scheme). Remediation: Evaluate trade-off between discarding a scheme/using other schemes and changing VisIt infrastructure.

Task A5. Start: 05/2008. Duration: 1 month.

Risk: Insufficient resources. Likelihood: high. Impact: High. Remediation: Delay final deployment.

Work Targets Next 12–24 Months

Once deployed and functional for our primary stakeholder (APDEC), we will apply this work to other application areas that use AMR.

12.7 Software Engineering

12.7.1 SCIRun Library for Volume Rendering (SLIVR)

Team

- Utah: M. Cole (Team Co-Leader) and J. Stratton
- LLNL: B. Whitlock and H. Childs (Team Co-Leader)

Stakeholder(s)

The primary focus of this project is software engineering to produce a "library version" of the high quality volume rendering technology in SCIRun for use in other areas, notably VisIt. The initial motivation for this work stems from stakeholder comments expressing a desire for better volume rendering in VisIt.

We aim to add 2d transfer function capability to VisIt's volume rendering infrastructure, similar to that present in SCIRun. Our approach is to separate the volume rendering engine and supporting code from SCIRun and turn that code into a separate library, which we call "SLIVR."

Another desired outcome for VisIt is to have better compositing, even for 1d transfer functions. VisIt currently does not have shading for its ray casted volume rendering mode, which leads to often poor quality pictures.

There are two broad phases of this project. The first is the software engineering work needed to create a library version of the SCIRun volume rendering technology. Part of this engineering effort includes using the new library in VisIt. The second broad software phase to software engineering to support a parallel use model on both shared and distributed memory architectures, particularly in VisIt.

Customer/stakeholder:

The initial motivation for this work stems from comments from some stakeholders requesting high quality volume rendering in VisIt. In the longer term, all VACET stakeholders who want to use volume rendering will benefit from this new capability.

Stakeholder need(s).

High quality volume rendering, multidimensional transfer functions to create more insightful images of multidimensional data.

Subtasks/milestones/deliverables.

- **Task 1.** Fully regression test, and merge with the SCIRun main development branch.
- **Task 2.** Integrate SLIVR lib with VisIt.
- **Task 3.** Create a 2d transfer function editor front end in VisIt.
- **Task 4.** Develop algorithm for distributed GPU cluster rendering, based on existing code.
- **Task 5.** Test/bug fix.
- **Task 6.** Optimize algorithms for efficiency and speed, including parallelization.
- **Task 7.** Clean up SLIVR interface.
- **Task 8.** Provide a SLIVR wiki for documentation and bug reporting.

Task Dependencies. The above list is mostly serial.

Subtasks. Marty and Josh are handling the SCIRun side, Hank and Brad the VisIt side. They are collaborating on new algorithms in SLIVR. Specifically, Hank and Brad are responsible for tasks 3. Task 4 is a shared task. Marty is responsible for the other tasks, many of which will be ongoing for the life of SLIVR. Hank and Brad will provide feedback and bug reporting to help with the maintenance tasks.

Estimated duration, start date, completion date. Started Nov 07, target completion: Nov 08

Tasks 1 through 8 are serial in terms of dependency.

Subtasks: Marty is handling the SCIRun side, Hank the VisIt side. They are collaborating on new algorithms in SLIVR. Specifically, Hank is responsible for tasks 6-8, although he may take existing code from Marty for task 7. Marty is responsible for the rest of the tasks, although Hank will likely provide input on 12 and 13 (and potentially others).

Estimated duration:

Project began in November 2006 and is expected to be completed in November 2007.

Accomplishments

Tasks 1 and 2 have been completed. SLIVR is now used in the main SCIRun development repository. It is fully integrated with SCIRun. Brad has accomplished Task 2 (see Figure 20.)

We have created a test application within SLIVR to verify installation and provide example usage code for anyone integrating the library into another application (see Figure 21.)

Josh has upgraded SLIVR to use GLSL for its shader language, which makes the code more maintainable/extensible. He has also added two new widget types to the 2d transfer function widget set.

Task 3 has been discussed in the monthly SEG telecon. The example application shows the usage of the widget code, which should make completion of this task more easily accomplished.

Task 5 has been discussed and preliminarily designed, and face to face meeting on this topic has been planned.

Work targets for the next six months.

Tasks 3,4,5 and 9. The biggest risk at this point is probably Task 5. This design may require a number of optimization passes, which can be somewhat open ended time wise.

Risks

- This project depends on resources being available from both the SCI side and the VisIt side. Without both sides, many of the tasks cannot move forward.

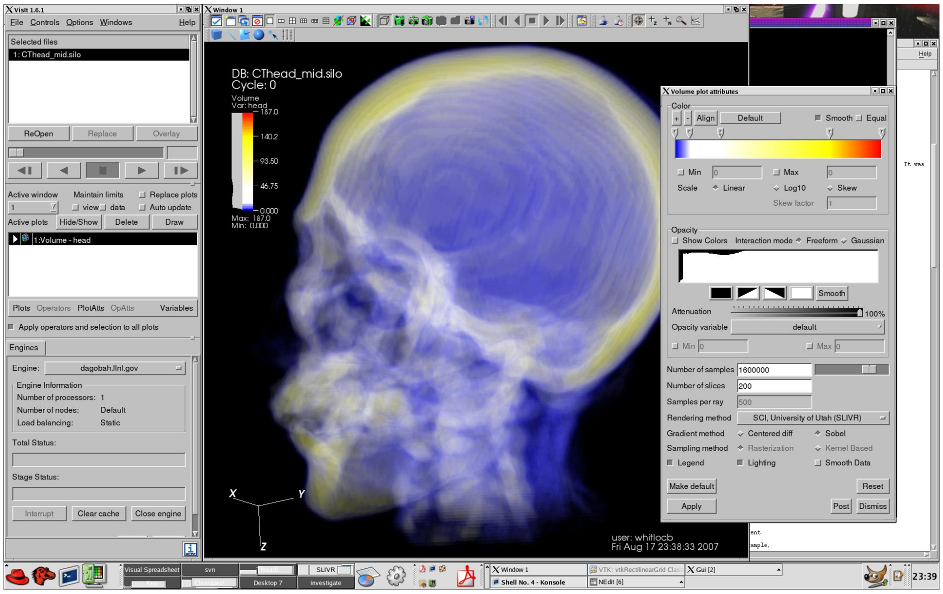

Figure 20: This image shows SLIVR being used in VisIt for volume rendering.

- Task 4 (distributed GPU rendering) is difficult and implementing a high performing solution is not a matter of straightforward software engineering.

Next 12-24 months. Task 5 will likely extend into this period.

12.7.2 VisIt Collaborative Development Environment and Infrastructure

Team

- Jeremy Meredith, ORNL (Team Leader) and Sean Ahern, ORNL.
- Hank Childs, LLNL

Stakeholder(s): The stakeholder is VACET itself. We are depending on VisIt to be a deployment vehicle and need a "hub" for the project. Although VACET could not undertake this effort by itself, it can with leverage from the SciDAC Outreach Center.

Stakeholder Need(s):

The high level need is to have the infrastructure available to manage a large scale software project. Although the VisIt project has previously made some of these investments, they are only accessible to LLNL, meaning that they are not aligned with the future of the project and need to be discarded and/or harvested. The VisIt project itself is highly supportive of migrating the software management infrastructure to the GForge server and will provide leverage for this effort.

Milestones/Deliverables/Objectives

Need A: Migrate the infrastructure for running the VisIt project to a "hub" accessible to VACET developers (and others):

- **Task A1.** Bug tracker.
- **Task A2.** Mailing lists.

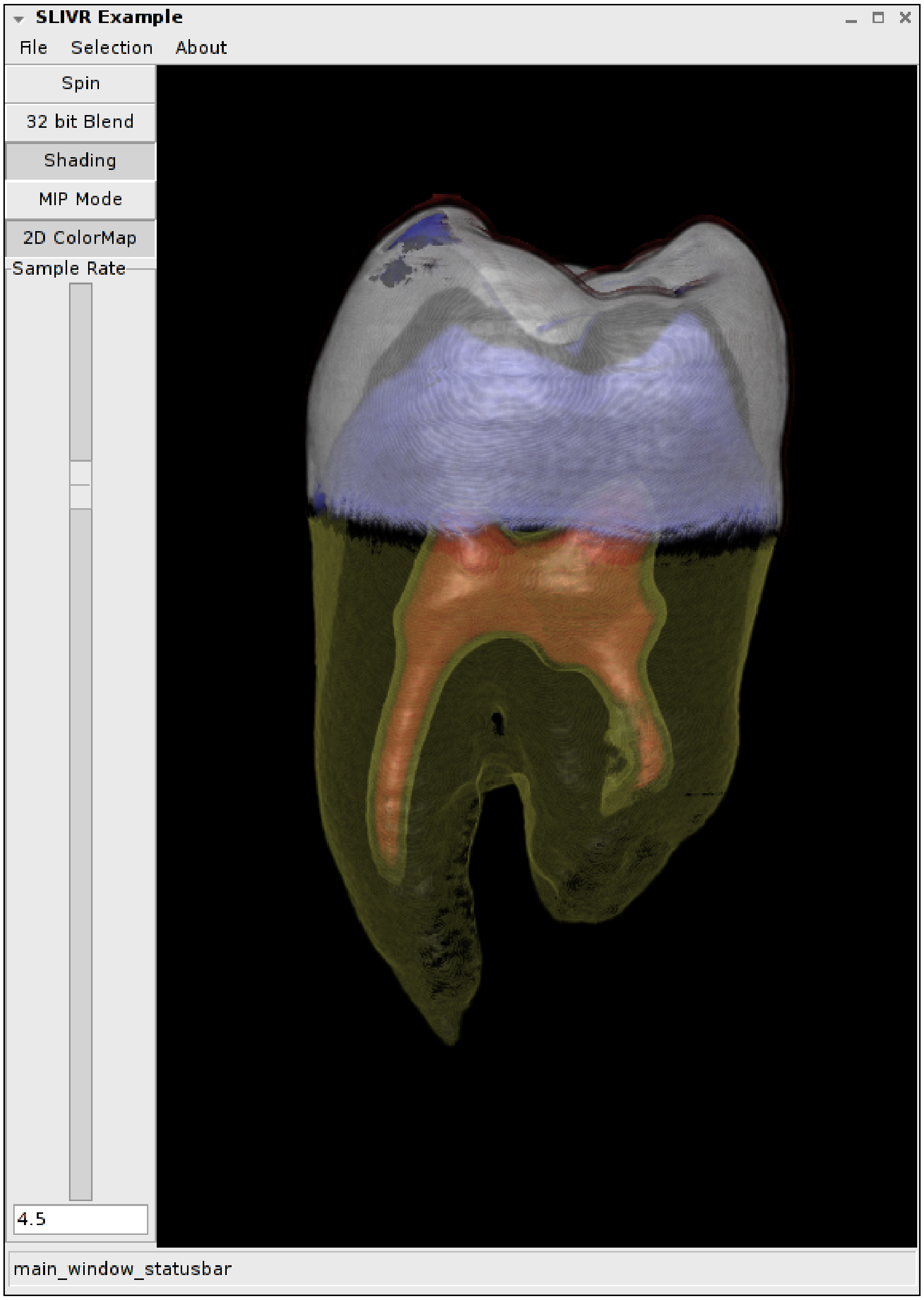

Figure 21: SLIVR test application using a 2D transfer function.

- **Task A3.** Wiki.
- **Task A4.** SVN access.

Dependencies

- Each task i depends on task $i - 1$.
- Task 4 – SVN Access – depends upon the completion of the project "VisIt Transition to Subversion" (see Section 12.7.3).

- All tasks depend upon the SciDAC Outreach Center. Specifically, these depend on their GForge installation, as this is the current implementation path.

Subtasks

- **Task A1.** Bug Tracker
  - **Task A1.1** Investigate needs of enhancement and defect tracking.
  - **Task A1.2** Investigate capabilities of GForge's built in issue tracker.
  - **Task A1.3** Create new tracker environment in GForge.
  - **Task A1.4** Export old tickets from existing ClearQuest tracker and import into GForge.
- **Task A2.** Mailing Lists
  - **Task A2.1.** Investigate needs for mailing lists.
  - **Task A2.2.** Investigate capabilities of GForge and the Outreach Center for mailing lists.
  - **Task A2.3.** Create new mailing lists utilizing the Outreach Center, archived at GForge.
- **Task A3.** Wiki
  - **Task A3.1.** Investigate needs for wiki documentation.
  - **Task A3.2.** Investigate capabilities of GForge's built in wiki plugin.
  - **Task A3.3.** Enable editing in GForge's wiki plugin for VisIt.
- **Task A4.** SVN Access (incl. anonymous)
  - **Task A4.1.** Export from existing SVN repository.
  - **Task A4.2.** Import into new GForge hosted repository.
  - **Task A4.3.** Confirm stability.
  - **Task A4.4.** Finish transition (export and shut down old repository, import into new repository).

Accomplishments

- Hank Childs and Gunther Weber met with Andrew Uselton of the SciDAC Outreach Center for a GForge tutorial and capabilities discussion on August 14th.
- Jeremy Meredith met with Andrew Uselton for a GForge tutorial and capabilities discussion on August 22nd.
- Hank, Jeremy, and Gunther met with the VisIt team to discuss how the GForge server would be used, and the requirements for various capabilities, on August 23rd. (This discussion led to the completion of A2.1 and A3.1.).
- Hank and Jeremy regularly provided input to David Skinner about the needs of VACET in order to ensure the GForge server would meet their needs when it was deployed.

Work Targets for Next Six Months

We intend to complete this project over the next six months. Note that there is a "go / no go" decision for each of these tasks. If the decision is "no go", then the dependent tasks will be re-thought.

We will definitely perform the investigative tasks: A1.1, A1.2, A2.2, and A3.2.

We will almost certainly perform: A1.3, A1.4, A2.3, A3.3, A4.1, A4.2, A4.3, and A4.4.

Risks

- Risk: GForge may not be capable of meeting some of our needs. Likelihood: medium. Remediate by: finding alternate solution. Impact: it depends, but likely medium.
- Risk: The SciDAC Outreach Center decides to no longer provide this service. Likelihood: low, especially since the majority of the cost has already been paid. Remediate by: either picking up support for GForge or abandoning this effort. Impact: high.

Work Targets Next 12–24 Months

We anticipate completion of this project in the next six months.

12.7.3 VisIt Transition to Subversion

Team:

- **VACET:** Hank Childs, LLNL (Team Lead).
- **Non-VACET:** The VisIt software team from LLNL..

Stakeholder.

There is no one single stakeholder for this project. It is primarily aimed at improving accessibility to the VisIt code base as well as providing better access/service to all who want to download, use and contribute to VisIt. Therefore, the stakeholder base is very broad. In addition, the ASC program is a beneficiary, and is providing personnel/effort in support of this project who will work with Childs and others on the VACET team.

Stakeholder Needs

Presently, there is no public access to the VisIt source code. This can be a problem if you want to run VisIt in parallel execution mode – the system must be compiled from source; there is no simple nor straightforward mechanism to provide binaries for parallel use due to the high degree of variability parallel execution environments. Typically, these are very site specific in implementation.

Related, there because there is no "public access" to the source code revision control system, it is very difficult for remote, distributed and non-LLNL VisIt developers to contribute to the project.

Milestones/Deliverables/Subtasks

- **Task 1.**Establish a repository at NERSC. Convert the existing VisIt repository to SVN. Get existing developers accounts on `svn.nersc.gov`. Completed 4/2007..
- **Task 2.** Make modifications to the code so that the code will build, run, and pass regression tests. Completed 4/2007.
- **Task 3.** Customize Subversion environment to ease developer transition, for example with merging.
- **Task 4.** Get nightly regression testing working using the `svn.nersc.gov` code base.
- **Task 5.** Team moves away from ClearCase at LLNL to `svn.nersc.gov` at NERSC.

Task Dependencies

- Task 2 depends on Task 1.
- Tasks 3 and 4 depend on Task 2.
- Task 5 depends on Tasks 3 and 4.

Estimated Duration

- Tasks 1 and 2: completed.
- Task 3: 3 days.
- Task 4: 2 days.
- Task 5: 3 days (for training).

Accomplishments

A1 and A2 were completed by 4/2007. The remaining tasks were completed in the reporting period ending 10/2007.

This project is now completed. The VisIt Subversion repository is available to all VisIt developers at `svn+ssh://svn.nersc.gov/svn/visit`. This repository was made available by the SciDAC Outreach Center, who have also provided Subversion consulting support. The repository has proven very stable. Further, a VisIt release was made using Subversion, stressing branch management, etc, and everything went very smoothly.

- Tasks 1 and 2: Completed.
- Task 3: in progress.

Next Six Months We expect to have completed all Tasks 1-5.

Risks

- Risk: Customization of Subversion for VisIt is difficult. Likelihood: low. Impact on project: low. Remediation: effort may double. Only Task 3 is affected.
- Risk: additional, unforeseen requirements arise before transition can be completed. Likelihood: medium – Childs has vetted a 9 page document on the conversion that has a 19-point checklist for doing the transition. Only the most prominent 5 bullets appear above as Tasks 1-5. Impact on project: medium. Remediation: The previously mentioned document was the best way to remediate this; it only serves to reduce the risk. Affected task(s): Potential addition of new task.

12.7.4 SSH Tunneling for VisIt

Team

- Jeremy Meredith, ORNL (Team Leader).
- The VACET Software Engineering Team, as needed.

Background

VisIt is a multi-process application that can run in a "client-server" mode where the local desktop machine contains a viewer and a remote machine runs the parallel computation engine. Unfortunately, network and site security policy add complexity that often defeats VisIt's ability to perform a direct connection between client and server. Connecting to a remote compute node is often impossible because these compute nodes are typically inaccessible except through a gateway/login node. Connecting back from the compute node to the user's desktop machine is thus typically preferable, but can still be hampered by any of the following:(1) machines with their host name set incorrectly (easily done in Windows); NAT firewalls routing to multiple machines; VPN connections; firewalls also blocking return access from the remote machine. These are common problems for all VACET customers and VisIt users. We have been working most closely with one of our stakeholders – fusion and accelerator researchers at Tech-X Corporation – to engineer a solution that permits effective remote and distributed VisIt use.

The best solution is to tap into the tunneling capabilities built into most SSH clients. These can open sockets at an arbitrary port on the destination host which get automatically forwarded to an arbitrary local port. There are still a number of complexities to address in any solution:

- The available remote ports are not known until after the connection is established, but one must specify the remote port numbers to use before making the connection.

- One solution is to guess remote ports randomly, but the probabilities of collision must be analyzed in detail.
- Tunneling multiple remote ports to each local port cuts down on collision probabilities, but is not portable.
- Remote listening ports can only be assumed to listen for connections from the same node; parallel compute engines thus need a workaround (known as the gateway node problem).
- Different platforms handle tunneling differently, so Linux/Windows/MacOS must all be tested and possibly fixed individually.

Stakeholder(s): The primary stakeholder for this effort has been the fusion and accelerator research staff at Tech-X Corporation. However, the solution will have a potentially much broader audience – anyone who uses VisIt in a remote/distributed configuration.

Stakeholder Need(s):

Need A. Allow VisIt network connections to work through multiple firewalls, VPN, routers, etc.

- **Task A1.** Outline and investigate potential solutions and feasibility.
- **Task A2.** Implement automatic SSH tunneling for VisIt socket connections.
  - **Task A2.1.** Implement basic 1:1 random-remote:fixed-local port tunneling for known UNIX target platforms.
  - **Task A2.2.** Enhance `qtssh` for Windows to support port tunneling.
  - **Task A2.3.** Account for additional issues on OSX.
  - **Task A2.4.** Add additional forward layer for clusters (i.e. compute nodes / gateway node tunneling).

Need B. Implement more robust tunneling – speculative future work.

- **Task B1.** Add one or more of the following methods for more robust tunneling:
  - **Task B1.1.** Forwarding specs enabled on demand, as needed.
  - **Task B1.2.** Implement a many-to-one remote:local port mapping to reduce chances of a collision.
  - **Task B1.3.** Forward all data through one SSH tunnel (i.e. through VCL).
  - **Task B1.4.** Use SSH Library (sshlib) instead of SSH command-line executable.

Accomplishments

All milestones for need A (A1, A2.1, A2.2, A2.3, A2.4) were accomplished. Tested and deployed to customers.

Work Targets for Next Six Months

None planned. However, speculatively need B may become a requirement depending on end-user feedback, necessitating milestone B1.

12.7.5 VisIt Database Options

Team

- Jeremy Meredith, ORNL (Team Leader).
- The VACET Software Engineering Team, as needed.

Background

When opening or writing a file, there is currently no way to specify options to the file format plugin. This can be necessary if the file has an ambiguous interpretation that a user might clarify, if the reader needs extra information not in the file, or if the reader can add capabilities when requested by the user. Some more specific examples:

- Reading a file containing multiple chunks: are these chunks spatial domains or a time sequence?
- A file contains multiple variables: which of these correspond to the X/Y/Z coordinates?
- A raw binary brick file reader is invoked: what are the dimensions?

When writing a file, some of these same options might appear, and more, such as how the coordinate arrays are to be labeled, should the file be written as a binary or ASCII file, and should multiple domains be split into multiple files or placed into one monolithic file?

For implementing, there are two main classes of ability to specify options: when the file format plugin to use is known, and when it is not known. The former occurs in a few specific instances, such as using the "File—Open" dialog box to choose a plugin explicitly, and when adding a plugin identifier to the "OpenDatabase()" python command in VisIt's command-line interface. In this case, the options can be immediately presented to the user. However, if the plugin is not known (because prerequisites such as default format and file extensions are not known in the VisIt front-end applications), there is significant additional complexity – either the options must be specified blindly by the user, or a multi-pass approach to opening must be taken, where a first attempt checks to see which plugin is the most likely match, options are presented, and then the plugin finishes opening the file. Tradeoffs between code infrastructure investment and interface ease-of-use are not trivial. (Exporting databases always requires an explicit plugin specification from the user and has no additional complexity.)

Stakeholder(s): John Cary, Tech-X Corporation, and the SciDAC Science Application "Framework Application for Core-Edge Transport Simulations."

Stakeholder Need(s) and Milestones/Objectives:

Task A1. Infrastructure for specifying database options.

- **Task A1.1.** Basic infrastructure for passing options when prerequisites are known.
- **Task A1.2.** Infrastructure for supporting options when prerequisites are not known.

Task A2. Graphical specification of options.

- **Task A2.1.** Create runtime-generated window for displaying and entering database options.
- **Task A2.2.** Allow users to specify options graphically when exporting databases.
- **Task A2.3.** Allow users to specify options graphically when opening files with user-specified plugin.
- **Task A2.4.** Allow users to specify options graphically when opening files with unspecified plugin.

Task A3. Python specification of options.

- **Task A3.1.** Create python routine to convert tuples into database options.
- **Task A3.2.** Allow users to specify options via python when exporting databases.

- **Task A3.3.** Allow users to specify options via python when opening files with user-specified plugin.
- **Task A3.4.** Allow users to specify options via python when opening files with unspecified plugin.

Dependencies

- A2.1 < A2.2, A2.3, A2.4
- A3.1 < A3.2, A3.3, A3.4
- A1.1 < A2.2, A3.2
- A1.1 < A2.3, A3.3
- A1.2 < A2.4, A3.4

Accomplishments

- A1.1: completed.
- A2.1: completed.
- A2.2: completed.
- A2.3: in progress.
- A3.1: completed.
- A3.2: completed.
- A3.3: in progress.

Work Targets for Next Six Months

- A1.2: Start Oct 2008; end Nov 2008; 5 days.
- A2.3: Start Aug 2008; end Oct 2008; 3 days.
- A2.4: Start Nov 2008; end Dec 2008; 4 days.
- A3.3: Start Aug 2008; end Oct 2008; 1 days.
- A3.4: Start Nov 2008; end Dec 2008; 2 days.

Work Targets Next 12–24 Months None; this project is targeted for completion in the next six months.

12.8 Turbulence

12.8.1 Quantitative Analysis of the Turbulent Mixing Layer in Hydrodynamic Instabilities

VACET Team

- V. Pascucci, LLNL (Team Leader).
- Peer-Timo Bremer (LLNL) core visualization techniques, data comparison, topology.
- Daniel Laney (LLNL) data management and streaming.
- Ajith Mascarenhas (LLNL) feature tracking.

Stakeholders

Andy Cook and William Cabot (LLNL), who are part of the SciDAC physics project entitled "Simulations of Turbulent Flows with Strong Shocks and Density Variations."

Stakeholder Needs and Accomplishments

The Livermore VACET group, developed for the first time feature based analysis of extremely high-resolution simulations of turbulent mixing created by Andy Cook and William Cabot, members of the SciDAC physics project for "Simulations of Turbulent Flows with Strong Shocks and Density Variations." The work focuses on Rayleigh-Taylor instabilities, which are created when a heavy fluid is placed above a light fluid and tiny vertical perturbations in the interface create a characteristic structure of rising bubbles and falling spikes. Rayleigh-Taylor instabilities have received much attention over the past half-century because of their importance in understanding many natural and man-made phenomena, ranging from the rate of formation of heavy elements in supernovae to the design of capsules for Inertial Confinement Fusion.

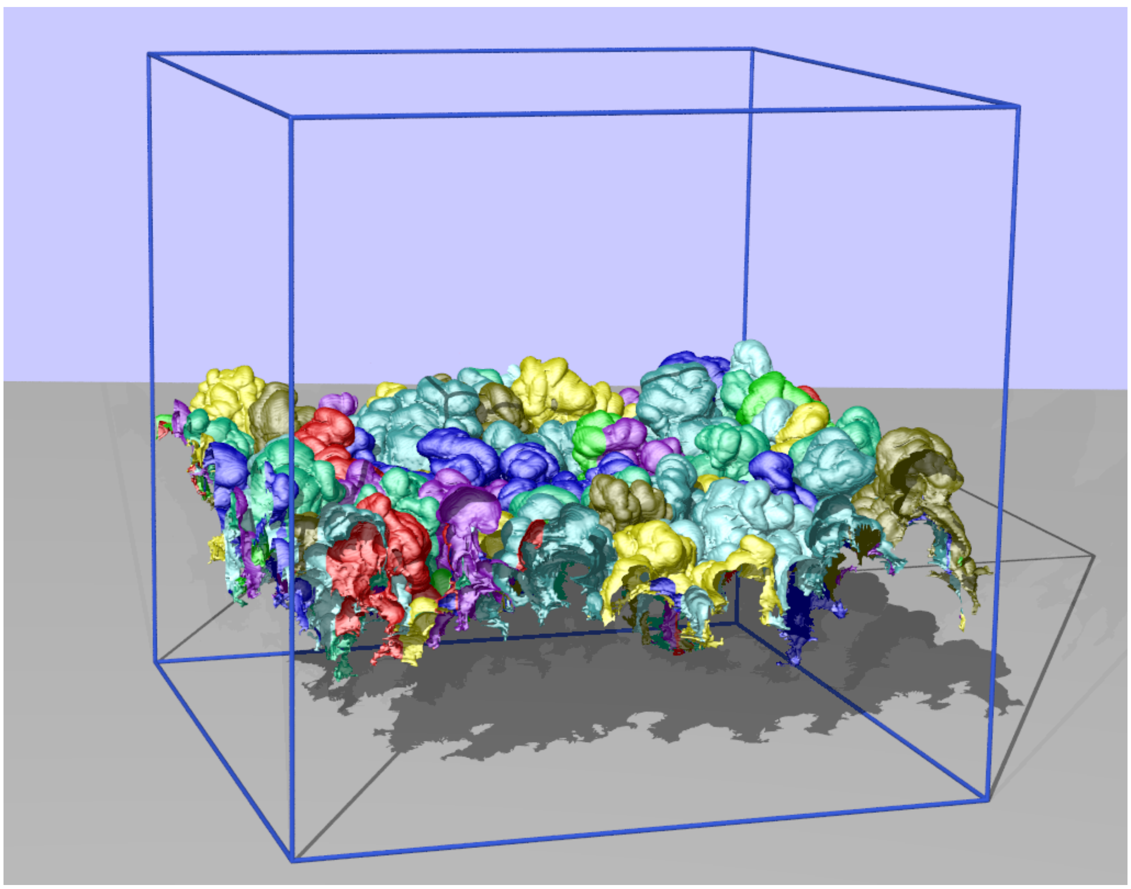

Figure 22: Topological segmentation of the bubbles emerging in the mixing interface of a Rayleigh-Taylor instability.

The VACET members at Livermore developed a novel approach to the analysis of the complex topology of the Rayleigh-Taylor mixing layer based on robust Morse theoretical techniques (Figure 22). This approach segments systematically the envelope of the mixing interface into bubble structures and represents them with a new multi-resolution model allowing for the first time a multi

scale quantitative analysis of the rate of mixing based on bubble count. The analysis highlighted and provided precise measures for four fundamental stages in the turbulent mixing process that previously scientists could only observe qualitatively, therefore enabling new insights and deeper understanding in this fundamental phenomenon (Figure 23).

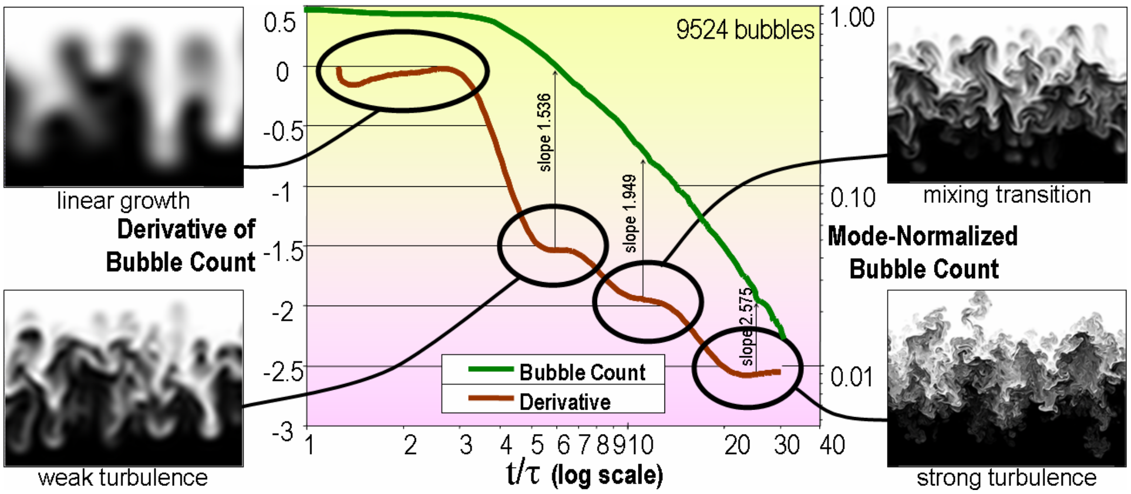

Figure 23: Quantitative analysis of a turbulent mixing simulation. The four flat regions of the derivative of the bubble count identify the four mixing stages (from early linear growth to late strong turbulence) and quantify the rate of mixing at each stage.

This work has been documented in a paper winner of the "best application paper" award at the IEEE visualization conference [33] and later presented at International Workshop on the Physics of Compressible Turbulent Mixing. Follow-up work allowed also for the first time direct comparison two simulations based on different physics models and run with different initial conditions; the first run with 1 billion nodes over 758 time steps, the second run with 27 billion nodes over 220 time steps. Although comparison by superposition of the two simulations could not yield any meaningful result, the topological approach allowed a multi scale feature based comparison highlighting fundamental similarities (Figure 24), which validated the lower resolution large eddy simulation (LDS) with respect to the higher resolution direct numerical simulation (DNS).

12.9 Comparative Visualization, Analytics and Analysis Technologies

12.9.1 Quantitative Analysis of the Impact of Micrometeoroids on Nanoporous Metals

Team

- Valerio Pascucci, LLNL (Team Leader)
- Attila Gyulassy (LLNL) 3D Morse-Smale complex construction.
- Peer-Timo Bremer (LLNL) core visualization techniques, data comparison, topology.

Stakeholder(s):

Material science simulation in LLNL, Primary contact Eduardo M. Bringa (LLNL).

Stakeholder Needs and Accomplishments

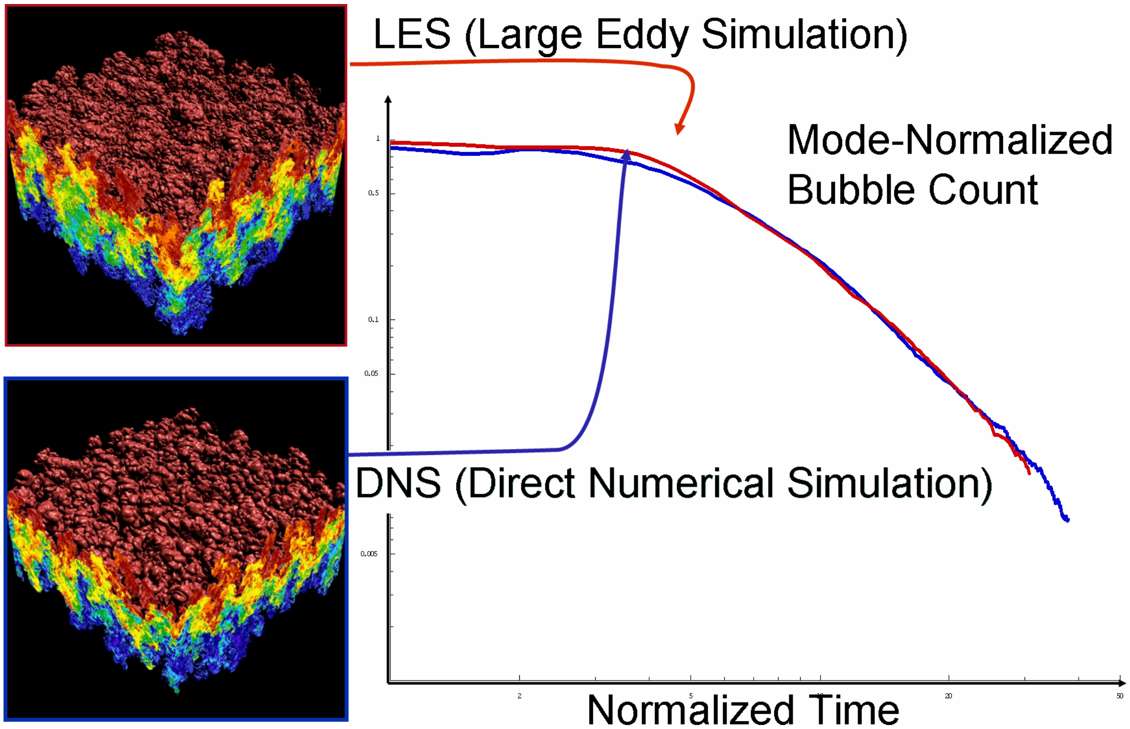

Figure 24: Feature based comparison of two Rayleigh-Taylor instability simulations. Although the simulations cannot compared by superimposition (different domain, different initial conditions), the feature-based analysis allows validating the Large Eddy Simulation with respect to the Direct Numerical Simulation.

There exists substantial interest in simulations of particle impact at multiple scales. Particles made of few-thousands of atoms are used to smooth or create nanoscale relief on a variety of surfaces. Millimeter-sized particles are often used to mimic impact of a variety of projectiles, from bullets to meteoroids. Recently, the Stardust mission[30] explored the craters left by micrometeoroids, reaching sub-micron sizes, in the frame of their comet dust catcher where the comet particles, moving at 6 km/s, were slowly decelerated, stored for recovery, and subsequently analyzed at laboratories worldwide. Understanding of such deceleration and storage is based on experiments and continuum-scale models of particle impact. Such models require equation of state and other thermodynamic input, and might fail at scales where atomistic effects would preclude continuum coarse graining.

To study the possible limitations of continuum models at the nanometer scale we have developed advanced topological tools for the robust quantitative analysis of the complex structures of porous media extracted from classical molecular dynamics simulations of the impact of a dense grain into low-density foam. Our new Morse theoretical approach analyzes the topology of the scalar field generated by the atomistic simulation are recover the most stable filament structures by filtration of unstable topological structures (see Figure 25).

Time tracking of the results provides answers to a number of fundamental questions including (i) quantification of loss in porosity of the material, (ii) description the modification of the density profile of the material, and (iii) quantification of the portion of the material that is affected by the

[30]http://stardust.jpl.nasa.gov/mission/index.html

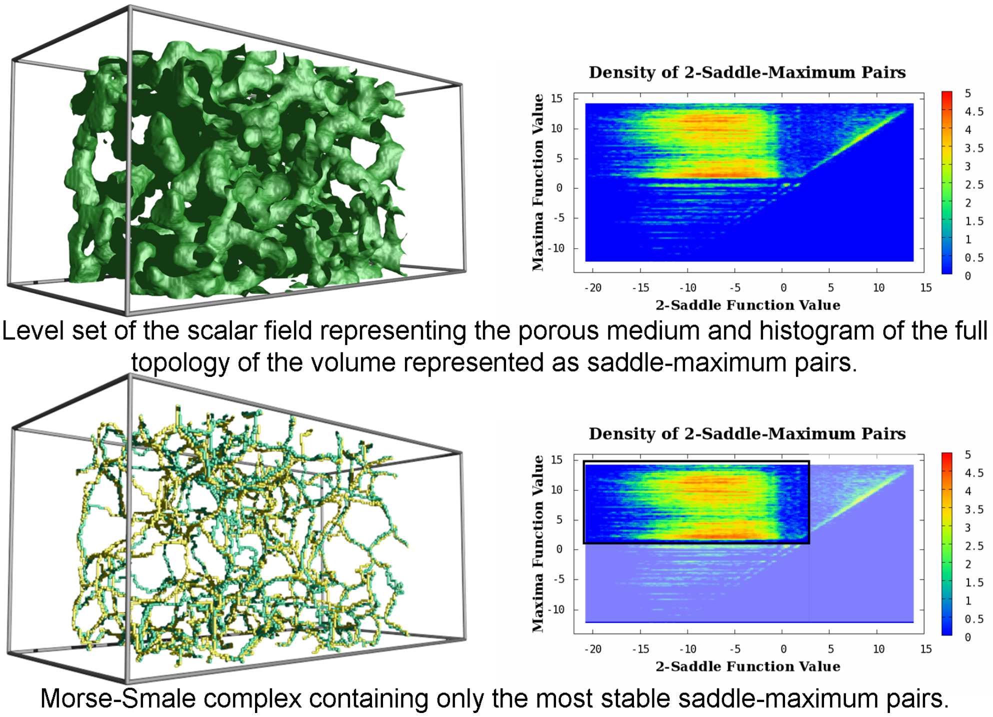

Figure 25: Geometric reconstruction of the porous medium from a filtered topology.

impact, and (iv) description of the structural modifications around the impact crater (see Figure 26). Overall we achieve for the first time a robust analysis of the structure of the structure of the porous medium as well as quantification of its dynamic behavior under intense structural changes.

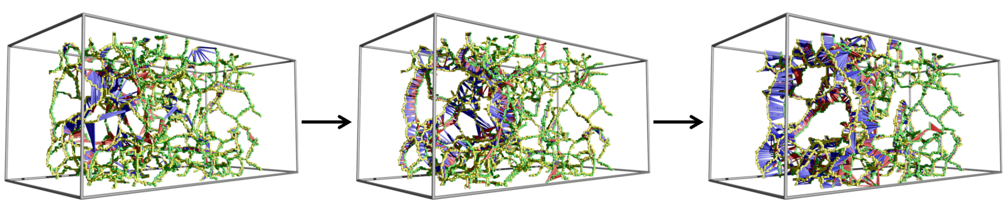

Figure 26: Deformation of the porous medium under collision with micrometeroids.

Figure 27 shows the evolution of the density profiles on the porous medium before and after the simulated impact event. Table 1 quantifies the degeneration of the structural strength of material by counting the total number of cycles and length of the network of filaments in the porous medium.

Publications

Attila G. Gyulassy, Mark A. Duchaineau, Vijay Natarajan, Valerio Pascucci, Eduardo M. Bringa, Andrew Higginbotham, and Bernd Hamann, Topologically Clean Distance Fields, Transactions on Visualization and Computer Graphics, IEEE Visualization 2007.

Work Targets for Next Six Months

While this project is largely finished, we plan on using the results of this work for the analysis of combustion simulations from John Bell at LBNL (see Section 12.4.3).

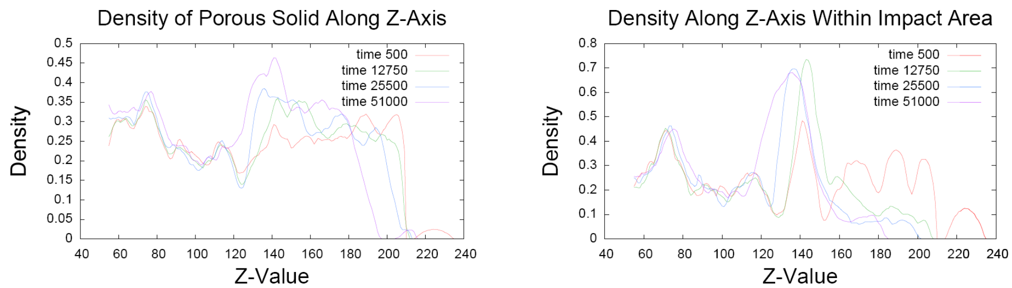

Figure 27: Density profiles of the porous medium at several time steps during the impact.

Metric	t=500	t=12750	t=25500	t=51000
Cycles	762	340	372	256
Total Length	34756	24316	23798	18912

Table 1: Topological analysis yields quantitative data useful for knowledge discovery. In this case, we have computed and report the number of cycles and path length of channels in the porous medium undergoing impact.

12.9.2 Multidimensional Visualization and Analytics – Parallel Coordinates

VACET Team: Sean Ahern, ORNL (Team Leader) and Jeremy Meredith, ORNL.

Stakeholder. For this work, there is no one specific stakeholder. Instead, this project focuses on designing and implementing infrastructure that would be applicable across a potentially large number of stakeholder projects.

While the original target was fusion, applicability extends to anyone with needs for high dimensional visualization. As the technique discards connectivity, it is one of the few suited as well for particles as for other mesh types.

Stakeholder Needs. Any project that needs to explore large, multidimensional/multivariate datasets would benefit from this work. During R&D, we have been using particle-based datasets produced by the GTC code from the Fusion community.

Milestones/Deliverables

- **Task A.** Summary information for high number of variables on large data sets. **(Completed Fall 2007.)**
  1. Implement 2D pairwise binning for arbitrary variable sequences.
  2. Add bin plotting to existing Parallel Axis plot with contrast enhancement.
  3. Add support for many-domain data sets, including in parallel.
  4. Add user interface for runtime selection of parameters.
- **Task B.**Focus areas of detailed exploration based on 1 or more variable ranges. **(Completed Fall 2007.)**
  1. Integrate summary information rendering (need A) with existing functionality to render all data points individually.
  2. Integrate user interfaces for the all-data-point ("focus") and summary ("context") plotting.

- **Task C.** Scaling to Arbitrarily Large Datasets. **(Completed Fall 2007.)** Remove barriers to scaling of algorithm to data sets of any size. Tested by running with a sample S3D dataset, courtesy J. Chen, SNL-CA, consisting of 60 million zones, on a distributed memory cluster.
- **Task D.** Automated outlier detection (without requiring manual plotting focus).
  1. Add basic outlier detection and plotting.
  2. Update infrastructure for multipass approach needed for parallel/multidomain detection.
- **Task E.** Make interface and interaction more intuitive and usable.
  1. Migrate visualization window-based widgets to explicit native GUI controls.
  2. Create new window dimensionality supporting more intuitive navigation (e.g. dimension-independent zooming and panning).

Accomplishments

Completion of Tasks A and B reported in the Spring 2007 PMP.

During this period, we completed scaling studies and delivered to users in a production release of VisIt, including testing with very large datasets.

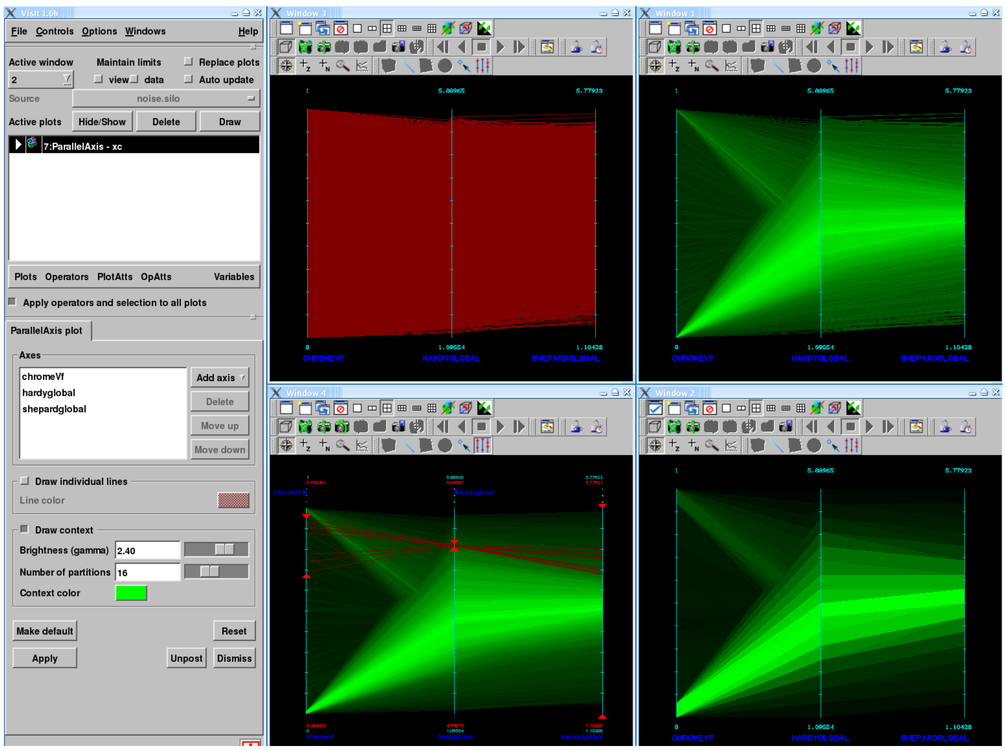

Figure 28: This image shows several examples of a parallel axis plot from one dataset. Clockwise from upper left: lines only (this part was not new; it existed before this project began and is shown mostly for contrast to the other images); context only with a high number of bins; context only with a low number of bins; context and focus of a select few data points.

Work Targets Next Six Months

None – project priorities have changed toward other tasks; the priority of this project was lowered to reflect this change. These changes have been necessitated largely due to the lack of resources in VACET.

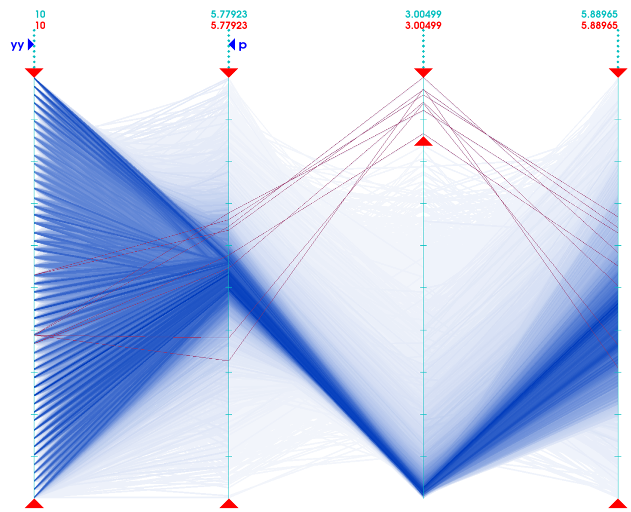

Figure 29: This image shows a more detailed example of the parallel axis plot.

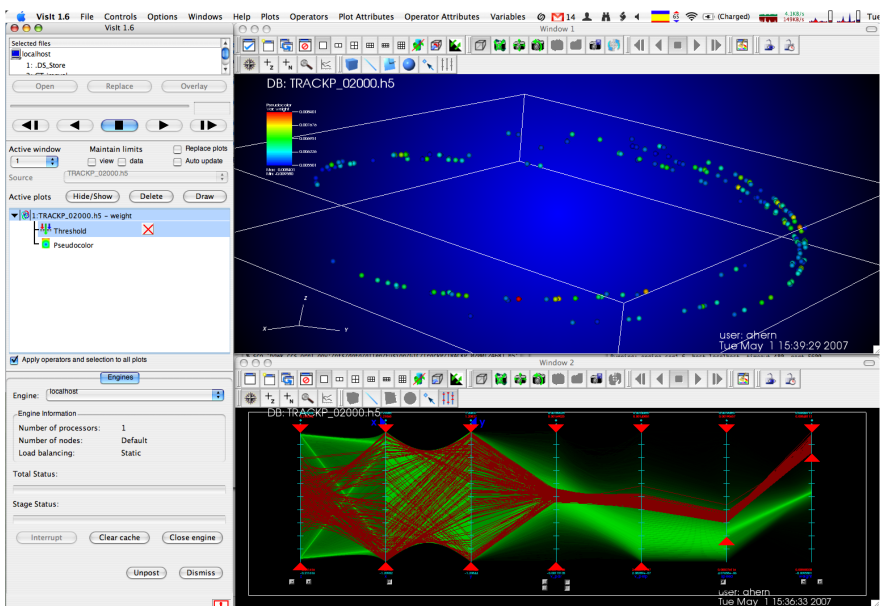

Figure 30: Here, we apply multidimensional filtering specified via the parallel coordinates interface in VisIt to select and display a reduced-size dataset of particles produced by the GTC code.

Work Targets Next 12-24 Months

Milestones for Tasks D (D1,D2) and E (E1,E2).

- Expected start date: April 2008.
- Expected duration: 2 months.

Risks that might delay project: insufficient funding.

12.9.3 Representative Particle Field Sampling

VACET Team: Sean Ahern, ORNL (Team Lead); George Ostrouchov, ORNL; Jeremy Meredith, ORNL.

Stakeholder(s): Stephane Ethier, PPPL of the SciDAC Center for Gyrokinetic Particle Simulations of Turbulent Transport in Burning Plasmas (W. Lee - PI).

Stakeholder Need(s):

It is not possible to fully browse a petabyte of data. To see 10% of a petabyte of data at the rate of 100 MB per second takes 35 8-hour days. At this rate, we are clearly exceeding the limits of of human bandwidth and patience. Although simulation brings with it the notion of unlimited transparency of simulated systems, as data sizes approach a petabyte, the difficulty of browsing such vast amounts reintroduces a great deal of opacity. There is a need for analysis methods that assist the human scientist in exploration of data by organizing the data and suggesting (or finding) interesting behavior. GTC particle simulations can easily produce petabyte data sets, although present output is usually limited to a random subset of the simulated particles. A particle selection capability is needed at two stages: in the simulation code itself to limit the amount of I/O, and further during visualization to provide interactive rendering speeds for the data that has been output.

A random selection of particles does not represent extremes well and over represents frequent behaviors. The essence of a "representative" sample is to present all "distinct" behaviors along with a measure of how many particles are represented by each behavior. Statistical methods that take into account behavior distributions can provide the stratification necessary for representative samples. Many clustering methods and sampling methods ranging from Latin square design to various fractional factorial designs are available in the R system for statistical computing[31]. These methods are needed for efficient exploration of high dimensional relationships with relatively small samples.

A major advantage of statistical particle selection over simple browsing or just taking the first so many particles is the availability and use of sampling distributions that provide uncertainty information for the samples. The sampling distributions are also the basis for automated selection of areas of interest. Interesting areas of dependence (or structure) can be discovered by comparing frequencies expected under an assumption of independence to actual frequencies.

The high level components needed to implement this capability include:

- A. Exploration of the problem and working relationships.
- B. Development of data-parallel runtime environment for R.
- C. Data-parallel HDF5 GTC particle data reader for R.
- D. Definition of particle attributes of interest.
- E. Scalable attribute density estimation methodology.
- F. Particle selection methods based on attribute densities and representation of sample uncertainty.
- G. Statistical methods to select particles of interest based on expected or unexpected behavior, etc.

Milestones and Deliverables

- A. A revision of the project wiki that reflects a better understanding of the problem.

[31] http://www.r-project.org/

- B. An evolving data-parallel GTC package for R (to be named RGTC).
- C. Data-parallel HDF5 particle reader for R added to RGTC.
- D. A collection of RGTC functions to calculate attributes of interest.
- E. (1) Histogram computation in RGTC (2) Density estimation in RGTC.
- F. (1) Latin square sampling (2) Fractional factorial sampling (3) Embedding of RGTC particle selection into SCIRun and VisIt.
- G. (1) Attribute interaction (dependence) search methods.

Dependencies

- Task A is ongoing and in parallel with the remaining tasks.
- Task B precedes all the remaining tasks and its revisions trigger revisions to the tasks.
- Task C follows and may be revisited with new data sets.
- Tasks D and E can be done in parallel.
- Tasks F and G follow and can be done in parallel.

Accomplishments

- Task A. The current revision (Fall 2007) of this page completes this task for this reporting period.
- Task B. A data-parallel runtime environment based on Rmpi[32] has been built for the hawk (ORNL) visualization cluster. The main difficulty was to provide an interactive system that retains complete readline and graphics interactivity of the R system. It remains to bundle it as a package.
- Task C. A data-parallel R reader for GTC particle files is completed based on the current version of the runtime environment. This will be bundled in the RGTC package. The reader performs best when files are stored on a parallel file system such as Lustre. Current GTC output consists of one file per time step and the reader distributes the time dimension across the nodes. Depending on the type of analysis performed, this may or may not be the optimal data organization. There is always a tradeoff between parallel analysis code complexity and data movement for reorganization.
- Task E. A data-parallel histogram R function, mpi.hist, is completed. The function computes local histograms in parallel and then merges the results. This works well with the current data organization. A few data reduction functions for computing particle attributes were also tested.

Work Plans for Next Six Months

- Task B. Add generality and "package" the runtime environment RGTC.
- Task E. Add density estimation methods to the histogram capability.
- Tasks F-G. Begin experimenting with Latin square sampling and interaction testing.

[32] http://cran.stat.ucla.edu/src/contrib/Descriptions/Rmpi.html

12.9.4 Variable Interactions in Query-Driven Visualization

VACET Team: Luke Gosink, UCD; John Anderson, UCD; E. Wes Bethel, LBNL; Kenneth I. Joy, UCD (Team Leader)

Stakeholder(s): John Bell and Marc Day, LBNL (Combustion Research).

Stakeholder Needs.
Query-driven methods are among the small subset of techniques in which we are able to address petascale data sets. However the challenges of efficient management of the data, and visualization of the results is only part of the problem. With our increasing capacity to generate larger and more complex data sets, there is a need for **scalable methods** that **provide insight** into the trends and behaviors of myriad of variables in a simulation. Our team is developing new methods by which coherent and meaningful visualizations can be constructed that convey information about the interactions between quantities in a query – vastly improving the scientist's ability to discover interactions between variables in multivariate simulations. We are currently collaborating with John Bell and his team at Lawrence Berkeley National Laboratory in applications of flame-front discovery to illustrate this technology.

Milestones/Deliverables The primary objective of this project is to integrate useful visualization techniques together with query-driven methods. Currently query-driven methods only provide the user with a "black box" that returns cells that satisfy a query. We wish to develop methods that allow interactions between variables in a query to be visualized.

We expect to integrate these techniques into VisIt and develop modifications for the method according to the needs of scientist stakeholders.

Accomplishments

- **Initial Prototype.** We have developed an initial prototype that allows multiple variables to be integrated into one visualization. Using Cumulative Distribution Functions and Correlation Maps, we can produce visualizations that illustrate the correlations between data against isosurfaces that are defined by the cumulative distribution functions.
- **Paper submission:** L. Gosink, J.C. Anderson, E.W. Bethel, and K.I. Joy, "Variable Interactions in Query-Driven Visualization" submitted to IEEE Visualization 2007.
- *October 2007 – Journal Paper.* L.J. Gosink, J.C. Anderson, and K.I. Joy, "Variable Interactions in Query Driven Visualization," IEEE Transactions on Visualization and Computer Graphics (Proceedings of IEEE Visualization 2007), 13(6), October 2007 [35].
- Project page: http://www.idav.ucdavis.edu/ joy/VACET/QueryDrivenVis.pdf.

Future Work
In the next 12-24 months, we expect to first integrate these techniques into the work activities of our scientist/stakeholders, allowing us to evaluate their usefulness. We will refine our methods as necessary to produce better techniques, depending on their needs. We also expect to develop new methods of visualization based upon the requirements for visualization of our stakeholders.

12.9.5 Equivalence Class Functions

Team
VACET: Hank Childs, LLNL; George Ostrouchov, ORNL; Sean Ahern, ORNL; Kenneth I. Joy, UC Davis (Team Leader)

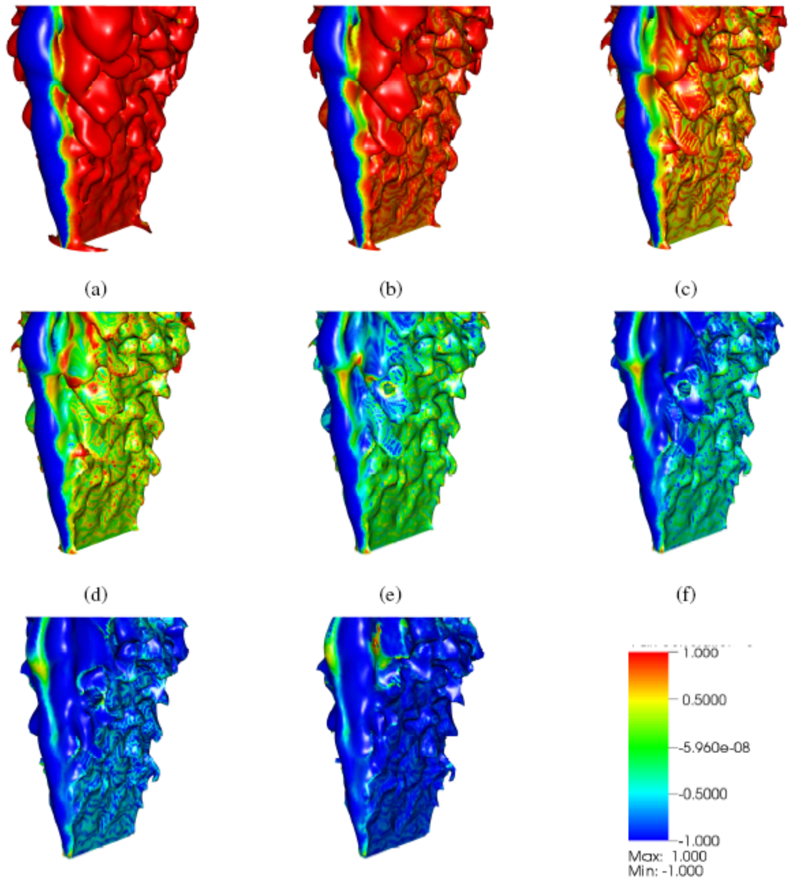

Figure 31: Proceeding from top to bottom and left to right, these images show isosurfaces of increasing values of temperature, which have been rendered through a correlation field constructed from H_2O and C_2H_4. As temperature values increase, we observe a full spectrum of color changes on each isosurface from red to blue indicating that the correlation between H_2O and C_2H_4 is not independent of temperature.

Non-VACET: Mark Miller, LLNL; John Clyne, NCAR.

Stakeholder(s). At the present time, this effort focuses on a need that is common across many different science stakeholders. Early versions of this work have been applied to combustion and turbulence research problems. The technique is generally applicable to a very broad range of visual data analysis problems.

Stakeholder Needs.

- Practical tools and algorithms for visualization and analysis of data at extreme scales.
- The ability to quickly and easily generate derived fields reflecting combinations of original data, statistics, or other user-defined operators.

Milestones/Deliverables/Objectives The use of statistical methods in scientific visualization is limited by the lack of scalable algorithms for statistical analysis. Our objective is to work with our

National Laboratory stakeholders to develop new scalable statistical techniques that can be used on petascale data sets. The additional analysis capability provided by these methods will greatly increase the data analysis tools available to researchers and scientists.

- **Task A1.** Develop the basic ECF Framework (Complete).
- **Task A2.** Integrate the basic framework into VisIt (Complete).
- **Task A3.** Develop new scalable statistical techniques that can be integrated into the ECF Framework.
- **Task A4.** Integrate these techniques into VisIt to develop new methods for analysis of scientific data sets.
- **Task A5.** Deploy in VisIt.

Dependencies

- A1 and A2 have been completed
- A3 requires completion of A1 and A2
- A4 requires partial completion of A3
- A5 requires partial completion of A4

Subtasks

Task A1. Work Team: Childs, Ahern, Ostrouchov, Joy

- Complete

Task A2. Work Team: Childs, Ahern, Ostrouchov, Joy

- Complete

Task A3. Work Team: Childs, Ahern, Ostrouchov, Joy

- Working with science stakeholders, identify new statistical operators that can be integrated into the ECF framework.

Task A4. Work Team: Childs, Ahern, Joy

- Introduce new ECF statistical operators into VisIt.

Task A5. Work Team: Childs, Joy, (future graduate student)

- Integration of specific operators into VisIt
- Testing and Debugging of implementation.

Accomplishments

VACET Researchers have developed a new class of derived quantities – Equivalence Class Functions (ECFs) – designed to greatly expand the ability of end users to explore and visualize data. These functions are defined over equivalence classes (i.e., groupings) of elements from the original mesh, and produce summary values for the classes as output. ECFs can be used in the visualization process to directly analyze data, or can be used to synthesize new derived quantities on an original mesh. The design of ECFs enable a parallel implementation that allows the use of these techniques on massive data sets that require parallel processing. When integrated and used in concert with standard visualization techniques, ECFs offer a powerful way to manipulate, visualize and analyze scientific data. Formal definition of ECFs and their efficient data-parallel implementation for data sets at the extreme scale will extend the application of many statistical methods to data set sizes they previously could not reach.

- Preliminary implementation of these concepts into VisIt, application to datasets from combustion and turbulence research.
- Publication: M. Miller, H. Childs, J. Clyne, G. Ostrouchov, S. Ahern, and K.I. Joy, "Frameworks for Visualization at the Extreme Scale," *Phys.: Conf. Ser. 78* (2007) 012035 (10pp)
- Project page: http://www.idav.ucdavis.edu/ joy/VACET/ECF.pdf.

Work Targets for Next Six Months
In the next 6 months, we expect to address Task A3, and work with our scientist stakeholders to develop new ECF-like functions that can be used to perform statistical analysis of large-scale data sets. This work is currently in progress.

Work Targets for Next 12-24 Months
In the next 12-24 months, we expect to address Task A4, implementing new functionality into VisIt which reflects our statistical ECFs.

12.9.6 Comparative Visualization and Analysis

Project Team:
VACET: Hank Childs, LLNL; Sean Ahern, ORNL; Jeremy Meredith, ORNL; Kenneth I. Joy, UC Davis (Team Leader)
Non-VACET: John Clyne, NCAR; Mark Miller, LLNL.

Stakeholder(s): This project will impact a potentially large audience – those who need the ability to perform comparisons on complex data sets. VACET has received requests from virtually all stakeholders for the ability to perform comparative analytics.

Stakeholder Need(s):
We concentrate on a data-level comparative visualization system that is very general in nature – including comparisons in physical space, comparisons in logical space, comparisons across symmetry conditions, and comparisons to analytical functions. Where many previous comparative visualization efforts have focused on visual comparisons, or A-B comparisons, our system is able to compare many related simulations in a single analysis, allowing for novel visualizations of ensembles of simulations or time-varying data.

Milestones/Deliverables/Objectives
Our objectives are to produce new data-level comparison techniques for scientific visualization. These techniques will be an enhancement of the data-level A-B comparison techniques, published in the past. We expect to develop a number of new comparison techniques and integrate them into VACET.

- **Task A1.** Develop basic framework for comparison through derived quantities. **(Completed)**
- **Task A2.** Develop cross mesh field evaluation techniques. **Completed**
- **Task A3.** Develop comparison methods for time-varying data sets that allow comparisons of the complete data set. **In Progress**
- **Task A4.** Develop comparison methods for parameter studies, that allow comparisons of the complete parameterized data set. **In Progress**
- **Task A5.** Integration of these techniques into VisIt **Partially Complete**

Dependencies

- A1 and A2 are somewhat independent and can be completed separately
- A3 requires A1 and A2
- A4 requires A1 and A2
- A5 requires A3 and A4

Subtasks

Task A1 – is complete.
Task A2 – is complete.
Task A3 Work team: H. Childs, K. Joy

- Develop comparative methods that can be used on complete time-varying data sets.
- Develop the tools (derived functions) that can be used to support these tasks.

Task A4 Work team: H. Childs, K. Joy

- Develop comparative methods that can be used on complete time-varying data sets.
- Develop the tools (derived functions) that can be used to support these tasks.

Task A5 – is partially complete.

Accomplishments

- Initial prototype implemented in VisIt. This prototype used for algorithmic testing and evaluation with a variety of different stakeholder datasets.
- H. Childs, S. Ahern, J. Meredith, M. Miller, and K.I. Joy, "Comparative Visualization Using Cross-Mesh Field Evaluations and Derived Quantities," *IEEE Transactions on Visualization and Computer Graphics*, submitted.
- Project Page: http://www.idav.ucdavis.edu/ joy/VACET/ComparativeVis.pdf.

Work Targets for the next Six Months

Tasks A3 and A4. Work in progress, duration six months.
Publish We will publish our work on comparative visualization of time-varying data sets.

12.9.7 Construction and Visualization of Function Fields

Team:

VACET: John Anderson, UCD; Luke Gosink, UCD; Kenneth I. Joy, UCD (Team Leader).
Non-VACET: Mark A. Duchaineau, LLNL.

Stakeholder(s):

Mark A. Duchaineau, LLNL (initial). As these techniques mature, we expect them to be applicable to radiation transport problems like those that are a part of S. Woosley's Computational Astrophysics Consortium project.

Stakeholder Need(s):

Stakeholders in this area require new techniques to visualize function fields. Radiation transport problems collect data in "bins," representing the data as histograms at each data point. Most of visualization (including the methods supplied by VisIt) is based upon the display of scalar fields, and these methods cannot be lifted to develop similar methods for function fields. Our objective is

to develop, together with our science stakeholders, a number of visualization techniques that allow the exploration of function fields.

Objectives/Milestones/Deliverables

Function fields map points in $n-$dimensional Euclidean space to one-dimensional scalar functions. These functions arise in neutron transport applications, hyper-spectral imaging applications, and air-pollution control modeling, among others. Our team is developing new methods for the exploration of such data. In general, they have developed probe-based methods that are used to derive scalar fields that show the similarity structure of function fields. Multiple probe methods allow more information to be seen and provide users with greater visualization flexibility over time and space.

- **Task A1.** Develop techniques that display features of function fields by utilizing metrics in "function space." **(Complete, Publication)**
- **Task A2.** Develop probe-based techniques that display features of function fields by allowing users to "influence" function-space metrics and better examine the features of function spaces. **(Complete, Publication)**
- **Task A3.** Integrate these techniques with the methods of our stakeholders.
- **Task A4.** Integrate these techniques into VisIt.

Dependencies

- A2 depends on A1.
- A3 depends on A1 and A2.
- A4 depends on A3.

Accomplishments and Future Work

Work in this area pre-dates VACET to July 2005. Together with the scientists that study the neutron transport problems (also called "energy-bin problems"), we have developed projects that display function fields. This work has resulted in one initial publication (EuroVis 2007). The multiprobe methods were developed with VACET support and will be published in IEEE Transactions on Visualization and Computer Graphics (the best journal in the field).

Work on this project stalled until J. Anderson received a Q-Clearance. Without the clearance, we could not adequately communicate with our LLNL stakeholders. This clearance was granted in September of 2007.

- Publication: J.C. Anderson, L.J. Gosink, M.A. Duchaineau, and K.I. Joy, "Feature Identification and Extraction in Function Fields," Proceedings of the European Visualization Conference (EuroVis 2007), May 2007, 195-201.
- J.C. Anderson, L.J. Gosink, M.A. Duchaineau, and K.I. Joy, "Exploration of Function Fields using Multiple Probes," IEEE Transactions on Visualization and Computer Graphics , to appear.
- Project page: http://www.idav.ucdavis.edu/ joy/VACET/FunctionFields.pdf.

Work Targets for the Next Six Months

We will concentrate on A3, which will integrate our current methods into the discovery of information from radiation transport data sets.

12.10 Visualization Technologies

12.10.1 3D LIC: Vector Field Visualization

This project focuses on developing and deploying a technique known as Line Integral Convolution (LIC), a well-known approach for vector field visualization. Prior to work on this project, neither of our deployment applications (VisIt, SCIRun) have this capability for visualizing vector fields. Our aim is to provide stakeholders with better vector field visualization capabilities in the production visual data analysis infrastructure.

In addition to production deployment, we aim to achieve several objectives that are novel in nature. First, our design and implementation will be modular in nature so that the LIC field, which is essentially an image in 1, 2 or 3 dimensions that represents a steady-state flow field, can be used to texture any arbitrary geometric primitives. Second, as a texture, rendering performance is accelerated by the hardware texture mapping capability ubiquitous on GPUs. Third, our design and implementation approach aims to provide support for use with AMR data, something that is completely novel.

VACET Team Members:

Gunther Weber, LBNL (Team Leader); Louis Feng, UCD.

Customer/stakeholder:

This new capability was first requested by John Bell early in the needs assessment process. The new capability has the potential to have a positive impact on a broad scientific audience.

Stakeholder need(s).

State-of-the-art vector field visualization capability, particularly for use with AMR-based datasets. The stakeholder requested the ability to see a vector field of one variable, e.g., momentum, mapped onto an isosurface of another variable, e.g., temperature.

Subtasks/milestones/deliverables.

Task 1. Implement 3D LIC for a single, uniform resolution mesh.
Task 2. Add infrastructure to VisIt for combining 3D LIC with contour plots (e.g., isosurfaces).
Task 3. Add mapper to VisIt that renders isosurfaces with the LIC texture.
Task 4. Test by resampling AMR grids to a single resolution.
Task 5. Extend to full AMR hierarchies.

Task Dependencies

Tasks 1 and 3 are independent. Task 4 requires completion of Tasks 1 through 3. Task 5 requires completion of Tasks 1 through 3.

Estimated duration:

Tasks 1 through 3 are in progress with an estimated completion date of the end of July 2007.
Task 4 is expected to be completed by the end of August 2007.
Task 5 duration is unknown at this time.

Accomplishments

During the past six months, Weber now has a working prototype implementation in VisIt. Weber took over project, performed a major clean-up of source code (previously started by a graduate student), performed a substantial amount of debugging.

Presently, we are almost finished milestones A2 and A3. Implementation of mapper both in hardware and software GL (for scalable rendering). Basic VisIt infrastructure work to complete milestone A2.

Set-up of VisIt build environment at LBNL.

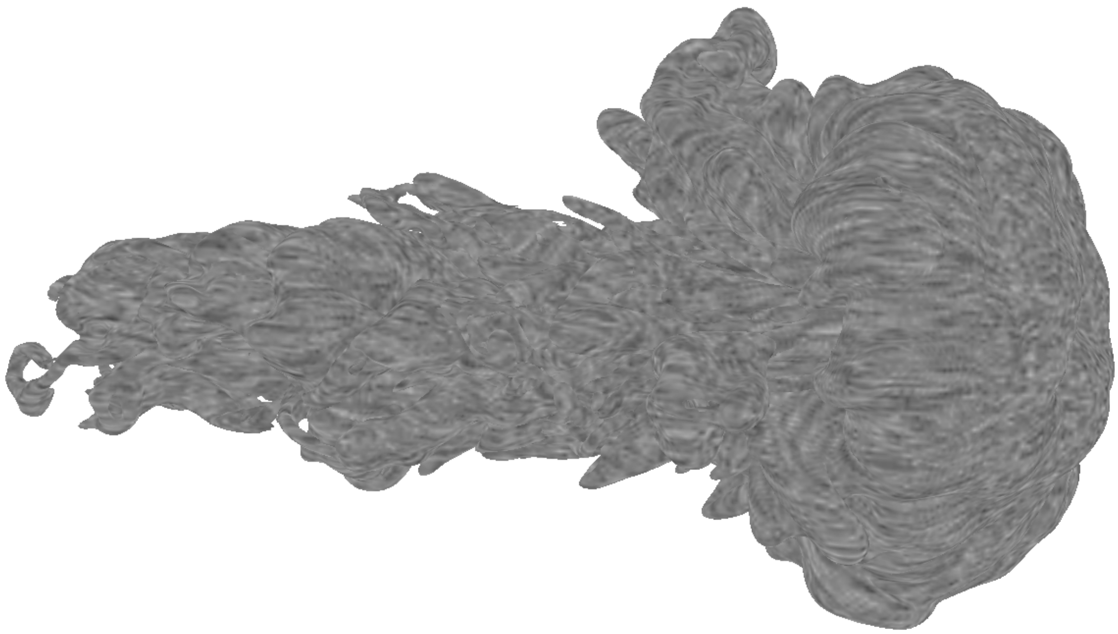

Figure 32: This example image shows the results of our prototype hardware accelerated 3D Line-Integral Convolution visualization technique. Here, the velocity field is depicted using 3D LIC on an isodensity surface. Our implementation is sufficiently general so that the 3D LIC field can be mapped onto any surface-based visualization technique, including isocontours, orthogonal or arbitrary slices, and even streamlines.

Next Six Months.
Re-evaluate customer needs and complete A4 and A5, or defer work.

Risks

- Potential reduced customer interest. Likelihood: medium. Impact: medium. Remediation: defer work until this topic is more of a customer priority.
- Inadequate personnel resources. Likelihood: medium. Impact: high – would not be able to make progress on the project due to inadequate staff resources. Remediation: defer work to the future.

Next 12-24 months.
Solidify code base to production quality and integrate into VisIt for a production release.

12.10.2 Unsteady Vector Field Visualization

Our research has led to the creation of a state of the art technique for the interactive visualization of unsteady flows on geometries of arbitrary complexity and topology. For that purpose we employ a dedicated atlas-based parameterization scheme that lends itself to visualization procedures involving global particle advection over the considered surface. The corresponding computation is mapped to a flat parameter space which is encoded on the texture memory of the graphics hardware to

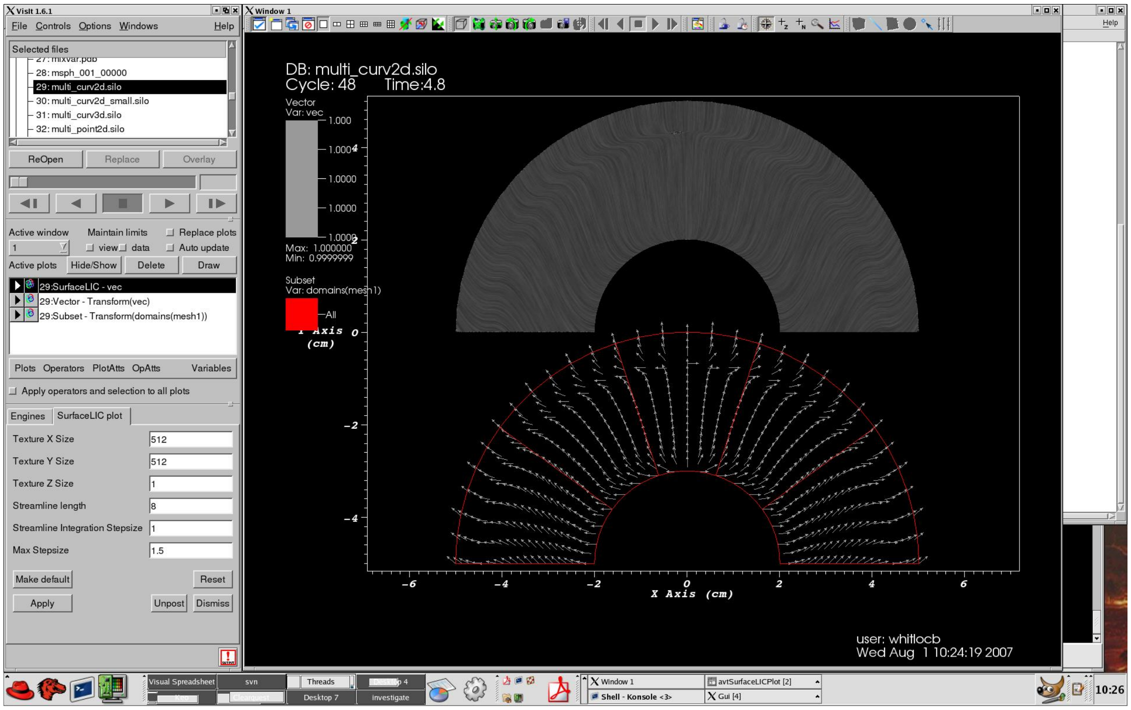

Figure 33: This image shows a screenshot of the prototype LIC interface in VisIt.

achieve interactive performance. Thanks to these critical features, our framework permits to easily extend most existing dense flow visualization methods to curved surfaces in 3D.

VACET Team: Guo-Shi Li, Utah; Xavier Tricoche, Utah; Chuck Hansen, Utah(Team Leader).

Stakeholder: Any scientific application within the scope of VACET where large 3D vector fields can be analyzed through their restriction to relevant surfaces embedded in the volume. Among them we find Combustion and Fusion.

Milestones/deliverables

- **Task 1.** Parameterize polygonal surfaces of moderate complexity in a way that supports particle advection in computational space.
- **Task 2.** Map corresponding data representation to the GPU.
- **Task 3.** Integrate existing dense flow visualization methods in the framework.
- **Task 4.** Extend the framework to surfaces of arbitrary complexity by improving mesh segmentation and parameterization aspects of the algorithm.

Start date: October 1, 2006.
Expected completion date: September 30, 2007.

Accomplishments

- Tasks 1 through 3 completed April 2007.
- Tasks 4 completed October 2007.
- Publication: C. Garth, F. Gerhardt, X. Tricoche, H. Hagen. "Efficient Computation and Visualization of Coherent Structures in Fluid Flow Applications," In Proceeding of IEEE Visualization 2007, pp. (accepted). 2007.

- Paper submission: G.-S. Li, X. Tricoche, D. Weiskopf, C. Hansen, Flow Charts: Visualization of Vector Fields on Arbitrary Surfaces. Submitted to IEEE Visualization 2007.

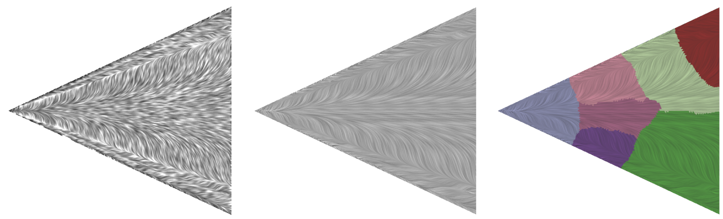

Figure 34: Above figure shows for a delta wing the mesh segmentation used to create the Flow Charts (right image) along with the seamless textures that it generates. Left image corresponds to a dense texture- based flow visualization method called UFAC (Unsteady Flow Advection Convolution, Weiskopf et al., IEEE Visualization 2003), while the middle image corresponds to GPUFLIC (GPU-based Unsteady Flow LIC, Li et al., Eurovis 2006).

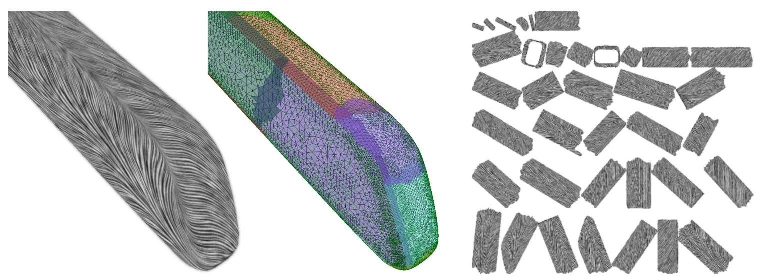

Figure 35: Above image shows similar results for an high-speed train. Left image corresponds to a GPUFLIC representation of the shear stress vector field on the surface of the train. Middle image shows a color coding of the segmentation superimposed on the mesh. Right image shows the packing of the individual pieces of the mesh in texture memory.

Future Plans

- Introduce analysis methods (topology, flow coherence) in the same framework to permit an integrated representation / analysis kind of interactive visualization.

12.10.3 Implicit Surfaces and GPU-based Volume Rendering

VACET Team: Chuck Hansen, Utah (Team Leader), Aaron Knoll and Josh Stratton, Utah.

Stakeholder(s) and Needs. These projects focus on fundamental, yet widely applicable technologies and represent more forward-looking aspects of the VACET portfolio. The results of these efforts would be applicable to virtually all VACET stakeholders desiring better/faster single or multi-field volume rendering capability.

Accomplishments

Fall/Spring 06/07: Aaron Knoll has been investigating implicit surface reconstruction in the context of surface reconstruction. This could be used in the ray-tracer. He has developed a SIMD-based interval arithmetic library for use on SSE (CPU) or G80 (GPU) platforms.

Josh Stratton's work has been reported under the SLIVR work (see Section 12.7.1. He has implemented a GLSL interface for SCIRun and SLIVR.

Work Plans Next Six Months

Josh Stratton will work on developing novel methods for multi-field volume visualization. The plan for late 2007 is to re-write the core SCIRun volume rendering and the associated SLIVR library (see Section 12.7.1. Currently, it is register-combiner code; he will generalize it to GLSL.

Obtain some sample datasets from VACET. Work on novel multi-field algorithms for pulling out interesting data relationships using volume rendering. Look at unstructured grid volume rendering.

12.10.4 Parallel, Out-of-Core Visualization Using Hybrid Resources

Team:

VACET: Brian Budge, UCD; Kenneth I. Joy, UCD (Team Leader)

Non-VACET: Tony Bernardin, UCD; Jeff Stuart, UCD ; Shubhabrata Sengupta, UCD; John Owens, UCD.

Stakeholder(s):

This research project is in collaboration with the SciDAC Institute for Ultrascale Visualization. We believe that most of the time-consuming algorithms for visualization can be implemented with an out-of-core methodology, using hybrid resources consisting of multiple GPUS. This research has the potential to impact a number of large-scale parallel applications through its novel use/management of resources.

Stakeholder Needs(s):

VACET researchers, in collaboration with researchers from the Institute for Ultrascale Visualization, are working to develop a new system that enables the rendering of very complex, out-of-core scenes with unbiased global illumination. The system facilitates the rendering of hundreds of millions of triangles, gigabytes of texture, and complex shading by forcing high levels of coherency. A hybrid processing paradigm is created, using both CPUs and GPUs, to exploit the data-parallel nature of rendering. The system has an out-of-core data management base layer designed to support a broad variety of rendering algorithms – including volume rendering for visualization purposes.

Objectives/Milestones/Deliverables

New software/system methods for high-quality visualization are required to address the complex petascale data sets developed by scientists at the National Laboratories. The objectives of this research effort is explore the feasibility of a highly-parallel hybrid approach to address this data.

Tasks

- **Task A1.** Develop system needs, including both software and hardware requirements for the prototype system **Complete.**
- **Task A2.** Purchase the prototype system. **Complete.**
- **Task A3.** Develop initial versions of the visualization management software. Develop initial prototype system **Complete.**
- **Task A4.** Evaluation of initial prototype, design/redesign phase. **Complete.**
- **Task A5.** Construction/purchase of additional equipment. **In Progress.**
- **Task A6.** Development of prototype system (version 2). **In Progress.**
- **Task A6.** System Evaluation. **Not Started.**

Subtasks

Task A1: Work Team: B. Budge, J. Owens, K. Joy – **Completed**

Task A2: Work Team: B. Budge, J. Owens, K. Joy – **Completed**

Task A3: Work Team: B. Budge, T. Bernadin, J. Stuart, S. Sengupta, J. Owens, K. Joy – **Completed**

Task A4: Work Team: B. Budge, J. Owens, K. Joy

- Evaluation of initial methodology.
- What are the additional requirements, both Software and Hardware to make this idea work?

Task A5: Work Team: B. Budge, J. Owens, K. Joy

- Determine new equipment to purchase, new software to develop.
- Purchase new equipment (if necessary).

Task A6: Work Team: B. Budge, T. Bernadin, J. Stuart, S. Sengupta, J. Owens, K. Joy

- Finalize implementation details.
- Construct system to support the final prototype.
- Final build of system.
- Final evaluation of system.

Accomplishments

- This project is complete through A3 – i.e., the initial prototype system has been developed. Evaluation testing is proceeding on the initial software infrastructure.
- A publication detailing this initial prototype has been submitted.
- We have added an additional controlling system (a master system, representing the remainder of our cluster as slaves). This will simplify the out-of-core management necessary to construct the prototype. This system has been ordered.
- Project page: http://www.idav.ucdavis.edu/ joy/VACET/Parallel.pdf.

Work Targets for the Next Six Months

Task A4. In the next six months, we expect to finish our analysis of the initial prototype and finalize the software infrastructure necessary to develop our final prototype.

Task A5. We expect to purchase a new controller for our cluster and implement this into the system.

Work Targets for the 12-24 Month Period

Task A6. We will implement our software architecture and build and evaluate our prototype system.

13 Risks

Our team has performed an evaluation of risk at both the Center level, as well as on a per project basis. We address Center-level risks below, while a risk analysis for each project has been performed and reported on a per-project basis in Section 12.

13.1 Principal Investigator

Scenario: One of the PIs leave the project. Likelihood: very low. Impact: high. Remediation: Due to the high degree of communication within the VACET center and, in particular, among the members of Executive Committee, the appropriate remediation is to have a backup PI from a lab and one from a University. Bethel and Johnson fulfill the roles of lead and backup PI, respectively.

13.2 Chief Software Architect

Scenario: The CSA leaves the project. Likelihood: very low. Impact: high. Remediation: we have designated Marty Cole (Utah) as the backup CSA. Both the CSA and backup CSA participate in monthly coordination meetings of the VACET software engineering team.

13.3 Current Budget Levels

Scenario: VACET's budget remains fixed at its present level for the entire five years. Likelihood: medium. Impact: Medium-High. When we engineered our proposal, it consisted of several broad but interlocking/interdependent technology areas that are absolutely required to achieve the objective of "PetaScale visualization." If any of them are eliminated, there is significant risk to VACET achieving its objectives. We felt it is in the best interest of our science stakeholders that we retain critical mass in some areas in order to maximize likelihood of impact. The alternative is to uniformly reduce effort across all technology areas. Technically speaking, that approach is viewed as being an even higher-risk – risk is spread to a greater number of project focus areas, thereby increasing overall project risk. Remediation tactics: (1) augment the VACET effort with SAP funding in the out years to pay for one-on-one stakeholder interactions, thereby freeing up some VACET "core funding" to be brought to bear on areas eliminated by the budget reduction; (2) lobby DOE to reinstate our funding based upon initial early successes with science stakeholders; (3) reduce the amount work from the original proposal; (4) look for opportunities to maximize leverage to achieve technical objectives.

Current Impact. The following list provides an overview of what we believe to be the current impact to our project of a 45 percent reduction in budget.

- **Eliminated the External Advisory Board.** This represents an estimated cost savings of about $25K per year – our original intent was to pay travel costs for EAB members to fly to a single location to conduct an annual review.
- **Reduced Effort in Multiresolution and Progressive Visualization and Rendering.** This effort offered promise for achieving dramatic scalability, but would have required a long lead time and a substantial investment, but would have put us into very good position to penetrate the challenges of Petascale and beyond visual data analysis.
- **Fewer customers.** At the present time, we do not have a concerted effort focusing on Accelerator modeling customers. However, the existing SciDAC Accelerator projects are concluding; the new Accelerator projects will begin later this year. At current budget levels, we are concerned about lack of manpower to dedicate to our Accelerator stakeholders when the new round of SciDAC funding begins. Related, we have been approached by several other Science Applications in the areas of Life Sciences, Materials, Chemistry, Groundwater and Climate; we have had to tell those potential customers that we do not have the ability to help them due to insufficient funding.

13.4 Computational Resources

Scenario: VACET, either collectively or in specific critical projects, is unable to obtain sufficient cycles or other computational resources to meet its stakeholders' requirements. Likelihood: low-medium. To date, VACET collaborators have received significant initial allocations from NERSC and ORNL. However, NERSC's interactive analysis platform, `davinci.nersc.gov`, has nowhere near the capabilities required to realistically achieve our stakeholders needs[33]. Impact: medium to high. Remediation: We would request assistance from DOE in obtaining the necessary resources. Failing that, we would prioritize resources in the direction of the highest priority projects and science stakeholders.

13.5 SciDAC Outreach Center

Scenario: The Outreach Center's scope does not include something like a GForge server that supports software downloads, bug tracking, email lists, collaborative content management, etc. Likelihood: medium. Impact: medium – the adverse impact to VACET is that VACET staff time would be required to set up the infrastructure to do software downloads, role-based collaborative document/content editing, and so forth. This effort will reduce the amount of time we have to accomplish our primary mission. Remediation: for software downloads, we could probably use SourceForge, but this approach is not viewed with favor. For collaborative document management and distribution, we may be able to use Google's 'Docs and Spreadsheets" (see `http://docs.google.com`), although this isn't an ideal solution. Recommendation: VACET recommends that the Outreach Center be immediately commissioned and activated according to the scope of its proposal.

May 2007: This risk seems to have been mitigated to a large degree – DOE has funded the SciDAC Outreach Center at a level commensurate need to provide the GForge-type service. The Outreach Center has recently hired a staff person to bring this new capability online. We await the formal launch of this new service. When fully operational, we will remove this Risk from our PMP.

14 Performance Measures

Generally speaking, the primary objective of *Performance Measures*, is to provide the means or a framework for evaluating the degree to which a given endeavor is meeting its objectives. As such, Performance Measures should be derived directly from a project's stated goals and objectives. For many endeavors, defining performance measures is straightforward and they are exclusively quantitative. Examples could include "Is profitability increasing?" In contrast, VACET's mission statement doesn't lend itself as well to traditional, quantitative and objective performance measures.

Earlier in Section 2, we listed a set of Goals and Objectives. Our performance measures, listed below, stem directly from those goals and are consistent with the overall SciDAC program mission.

It is well accepted that performance management of research activities is much different than that for other types of operational or service endeavors. For example, DOE's "Guidelines for Performance Measurement[34]" suggests that different frameworks for developing organizational performance measures are appropriate for different types of projects. The choice of one framework or

[33] davinci.nersc.gov is a 32P SGI Altix with 192GB of RAM, 24TB of scratch disk and no graphics hardware. We expect requiring a machine with on the order of 128 cores or more, 2TB of memory, 4GB/sec in I/O bandwidth, and the ability to achieve rendering rates on the order of 500M tris/second. To the best of our knowledge, there are no plans in place at NERSC to obtain such a platform

[34] http://www.orau.gov/pbm/documents/g1201-5.pdf

approach over another depends upon the type of project: a performance measurement approach appropriate for a construction project would not be appropriate for a research project.

Our approach to performance measurement reflects the diversity in scope and time horizon for the different goals in our project. Some performance measures are straightforward and can be obtained through normal operations. Others reflect accomplishments that are possible only over a long period of time and in the presence of favorable circumstances, many of which are outside our scope of control. This mixture reflects our research portfolio, which ranges from basic to applied, but with more emphasis on applied research aimed at meeting scientific stakeholder data understanding needs.

It should be noted that basic research doesn't lend itself well performance measurement in terms of "impact" since such impact often times doesn't occur until decades later. One such example is how the 1960's breakthrough of deciphering the genetic code has led to the identification of genes linked to illnesses such as breast and colon cancer, Huntington's and Alzheimer's disease, and the inception of gene therapy treatments. These impacts could not have been foreseen during the original research during the 1960s.

Finally, we will evaluate the effectiveness of these performance measures over the lifetime of the project. These are subject to change in future revisions of this Project Management Plan depending upon whether or not they prove effective and useful as performance measures.

14.1 Effective and Productive Stakeholder Relationships

Relationships with our customer/stakeholders is a crucial program element. Here, our aim is to work closely with stakeholder/customers to clearly define their scientific data understanding needs, translate those needs into actionable items, prioritize those needs taking into account both stakeholder input as well as VACET-wide considerations (e.g., limited budget, opportunity to reuse technology and effort across multiple stakeholders, unique opportunity for scientific or program-wide impact, etc.).

14.2 Customer Feedback

Customer/stakeholder evaluation is an important VACET performance measure. Here, customers are those who benefit or use, directly or indirectly, the products of VACET's R&D. Customer evaluation is the opinion of one or more customers about either (1) the extent to which VACET's program directly or indirectly benefits the customer; or (2) the extent to which the R&D is perceived as beneficial to the public.

14.3 Effective Production Deployment of Petascale-Capable Technology

As a CET, one of VACET's missions is to deploy petascale-capable visual data analysis software infrastructure at DOE's open computing facilities. In this context, "deploy" is interpreted broadly to include: (1) install the software at DOE's open computing facilities for use by anyone having access to those facilities; (2) run VACET software to process stakeholder datasets at those facilities; (3) provide assistance to the stakeholder in running VACET software at those facilities; (4) run or provide assistance in running VACET software on stakeholder datasets at a location of the stakeholder's choosing[35].

[35]This scenario – run at stakeholder facilities – reflects the reality that in some circumstances, it is more efficient to extract subsets of large datasets and move the subset to the stakeholder facility for more detailed analysis. In this case, VACET technology is still being applied to solve stakeholder problems, which meets the larger SciDAC program goal of using computational infrastructure to conduct scientific research.

14.4 Coordinated Interactions with Other SciDAC Technology Providers

We anticipate that there will be many opportunities to "combine forces" with other CETs, Institutes and Partnerships to bring to bear diverse technologies to solve stakeholder data understanding problems. This performance measure aims to reflect: (1) the degree to which VACET looks for such opportunities; (2) where appropriate, is able to capitalize upon such opportunities.

Seeking opportunity is expected to be part of our ongoing operations as we participate in community-wide meetings, attend workshops and meetings of our SciDAC sibling projects, have staff from other SciDAC projects attend VACET meetings, and so forth. This type of activity can be planned on an ongoing basis.

Capitalizing upon opportunity is a different matter, and cannot be planned in the same way. Capitalizing upon opportunity requires the confluence of several conditions, which may or may not arise at the right time and may or may not be within our control. Such conditions include: a given technology is sufficiently mature to adopt/apply to new applications in conjunction with VACET technologies; a stakeholder need arises that can be met by multiple technologies from diverse sources; the project teams from which diverse technologies originate have the labor resources to dedicate to a cross-team project. While we cannot predict when such opportunities will arise, we can plan to seek them out through ongoing operations.

14.5 Technical Excellence – Publications

Publications are an important vehicle for conveying research results to the broader scientific community. In addition, most technical publications undergo a process of peer review. The presence of publications in peer-reviewed forums is an important metric that shows our work is of high quality.

14.6 Outreach, Education and Service

Over the course of our project, we will also be engaged in several different forms out activities that fall into the broad category of "Outreach and Education." These activities include invited talks/presentations at workshops, conferences and so forth; participation in program-wide meetings, like the SciDAC 2007 program-wide meeting in Boston MA; workshops and tutorials.

A related set of activities falls under the general banner of "Service." These include things like: service as expert on a review committee, organizing and/or participating in planning workshops, service as a Program Committee member, organizer, or a reviewer for technical journals and conferences.

14.7 Train the Next Generation of Visualization and Analytics Leaders

An important aspect of VACET operations is providing the opportunity for junior staff to gain experience in leadership roles to prepare them to serve as future leaders.

15 Communication

15.1 VACET Internal Communication Vehicles

Since the VACET team is distributed both geographically and institutionally, we have come to rely heavily on three different forms of internal communication: email, telephone and wiki.

We have established two different email lists: one for the entire VACET team, the other for the VACET EC. These lists are hosted at LBNL and are quite active. Since we are using *Mailman* to

provide email list service, we also benefit from having access to browsable archives of email traffic on both lists.

We have established a standing, toll-free conference call line hosted at LBNL for use by team members. The objective is to make it as easy as possible for team members to quickly arrange conference calls. This project-wide service has proven very useful in ad-hoc meetings to address time-critical specific subjects. Examples of such meetings include: organization of a VACET workshop at the 2007 SciDAC program meeting; several meetings aimed at coordinating a VACET-wide response for Science Application Partnership proposals that are part of the Spring 2007 Fusion SciDAC call for proposals.

We have created a project-wide wiki that is hosted at SCI. The wiki has both public and authentication-required content areas. Its primary role so far is as a vehicle for posting and sharing documents, collaborative document development, individual stakeholder needs assessment and proposed work tasks, per-project management plans, and so forth. Related, SCI hosts a project-wide Subversion server that has proven very useful for collaborative, revision-control document and code development.

15.2 VACET Internal Communication Channels

The VACET management has adopted a position of encouraging frequent communication amongst all team members. This project-wide policy has proven to be highly beneficial in terms of achieving: (1) coordination of project-wide activities; (2) helping each team member to feel an active part of a larger, vibrant effort by knowing what everyone else is working on, project objectives, and so forth; (3) fostering an overall sense of unrestricted and open communication amongst all team members.

The VACET Executive Committee holds bi-weekly conference calls. The purpose of these calls is to discuss project management issues, including but not limited to: (1) stakeholder relations; (2) defining and prioritizing VACET technical objectives; (3) inter-institution technical interactions. Minutes of these calls are emailed to the entire team. In addition to the regularly scheduled bi-weekly calls, the EC periodically holds out-of-cycle calls to discuss time critical issues when needed.

The VACET Software Engineering Team holds monthly conference calls. Minutes of the calls are posted on the VACET wiki.

In addition to regular calls, these groups periodically hold ad-hoc meetings on an as-needed basis to address time-critical matters.

We have held two VACET-wide, all-hands meetings since the beginning of operations in October 2006. The first all-hands meeting was on 1 November 2006 in conjunction with IEEE Visualization 2006 in Baltimore, MD. The second was held January 31 and February 1, 2007 in Salt Lake City, Utah. The second meeting included representatives from our stakeholder population. More information about the second meeting is located on the VACET wiki[36].

15.3 Communication with Other SciDAC Projects

As indicated in previous sections, VACET team members have been very active in reaching out to stakeholder projects. Such outreach includes email, phone calls, and presence at their project meetings. Based upon the early success of such outreach, we expect such frequent and productive communication to continue in the future.

VACET was well represented at the February 2007 SciDAC project-wide meeting (The “Meet and Greet”) in Atlanta, GA. VACET will also be well represented at the upcoming SciDAC program meeting in Boston at the end of June 2007.

[36] http://www.sci.utah.edu/vacetwiki/index.php/2007_All_Hands_Meeting

As one of the guiding principles for VACET operations is free and open communication, we have a large amount of "surface area" with many different SciDAC projects: each VACET team member is an "ambassador" for the project. So far, this approach has been hugely successful in fostering broad communication with many different current and potential future stakeholders.

15.4 External Communication

Our strategy for outreach and communication to "the world" includes a broad array of activities: web, press, technical publications, and technical presentations.

VACET established a website at `www.vacet.org`. The website, while evolving, is intended to provide a comprehensive view of our Center's activities to a wide audience (science stakeholders, DOE program office, general public).

Press releases provide a valuable way to reach out to a potentially large audience that spans academia, government and industry. One example of such a press release is the November 3, 2006 HPCWire interview with VACET team member Bethel. These types of activities help not only VACET, but also SciDAC and DOE's technology portfolio as a whole.

VACET team members have made presentations in many forums since commencing operations. These are listed in Section 10.11.

VACET team members have a strong record of technical publications. We expect this trend to continue in the future, thereby maintaining VACET's leadership role in the visualization community. These are listed in Section 10.10.

References

[1] ASC VIEWS. `http://www.llnl.gov/asci/views`. Advanced Scienfic Computing VIEWS Program.

[2] Scientific Discovery through Advanced Computing. `http://www.osti.gov/scidac/SciDAC.pdf`. Office of Science U.S. Department of Energy.

[3] S. Ahern. Petascale Visual Data Analysis in a Production Computing Environment. 78, June 2007.

[4] J. C. Anderson, L. Gosink, M. A. Duchaineau, and K. I. Joy. Feature Identification and Extraction in Function Fields. In *EuroVis 2007*, pages 195–201, May 2007.

[5] R. Armstrong, D. Gannon, A. Geist, K. Keahey, S. Kohn, L. McInnes, S. Parker, and B. Smolinski. Toward a common component architecture for high-performance scientific computing. In *Proceedings of the 8th IEEE International Symposium on High Performance Distributed Computation*, Aug. 1999.

[6] Attila Gyulassy and Vijay Natarajan and Bernd Hamann and Valerio Pascucci. Efficient Computation of Morse-Smale Complexes for Three-dimensional Scalar Functions. 13(5), 2007.

[7] G. Bebis, R. Boyle, B. Parvin, D. Koracin, P. Remagnino, A. V. Nefian, M. Gopi, V. Pascucci, J. Zara, J. Molineros, H. Theisel, and T. Malzbender, editors. *Advances in Visual Computing, Second International Symposium, ISVC 2006, Lake Tahoe, NV, USA, November 6-8, 2006 Proceedings, Part I*, volume 4291 of *Lecture Notes in Computer Science*. Springer, 2006.

[8] G. Bebis, R. Boyle, B. Parvin, D. Koracin, P. Remagnino, A. V. Nefian, M. Gopi, V. Pascucci, J. Zara, J. Molineros, H. Theisel, and T. Malzbender, editors. *Advances in Visual Computing, Second International Symposium, ISVC 2006 Lake Tahoe, NV, USA, November 6-8, 2006. Proceedings, Part II*, volume 4292 of *Lecture Notes in Computer Science*. Springer, 2006.

[9] J. Bennett, V. Pascucci, and K. Joy. Genus oblivious cross parameterization: Robust topological management of intersurface maps. In *Proceedings of Pacific Graphics 2007*, 2007. To appear.

[10] F. Bertrand, R. Bramley, K. Damevski, D. Bernholdt, J. Kohl, J. Larson, and A. Sussman. Data redistribution and remote method invocation in parallel component architectures. In *Proceedings of The 19th International Parallel and Distributed Processing Symposium*, 2005. (Accepted, Best Paper Award).

[11] E. Bethel, C. Johnson, K. Joy, S. Ahern, V. Pascucci, H. Childs, J. Cohen, M. Duchaineau, B. Hamann, C. Hansen, D. Laney, P. Lindstrom, J. Meredith, G. Ostrouchov, S. Parker, C. Silva, A. Sanderson, and X. Tricoche. SciDAC Visualization and Analytics Center for Enabling Technology. 78, June 2007.

[12] P.-T. Bremer, E. M. Bringa, M. Duchaineau, A. G. Gyulassy, D. Laney, A. Mascarenhas, and V. Pascucci. Topological Feature Extraction and Tracking. 78, June 2007.

[13] P.-T. Bremer and V. Pascucci. *Topology-Based Methods In Visualization*, chapter A Practical Approach to Two-dimensional Scalar Topology. Springer Verlag, August 2007.

[14] S. P. Callahan, L. Bavoil, V. Pascucci, and C. T. Silva. Progressive volume rendering of large unstructured grids. *IEEE Transactions on Visualization and Computer Graphics*, 13(1):1307–1314, 2007.

[15] J. Chen, I. Yoon, and E. W. Bethel. Interactive, Internet Delivery of Visualization via Structured, Prerendered Multiresolution Imagery. Jan/Feb 2008.

[16] H. Childs. Architectural Challenges and Solutions for Petascale Postprocessing. 78, June 2007.

[17] H. Childs, E. Brugger, K. Bonnell, J. Meredith, M. Miller, B. Whitlock, and N. Max. A contract based system for large data visualization. In *Proceedings of IEEE Visualization 2005*, 2005.

[18] H. Childs and M. Miller. Beyond meat grinders: An analysis framework addressing the scale and complexity of large data sets (to appear). In *Proceedings of The 10th International Conference on High-Performance Computing*, 2006.

[19] K. Damevski and S. Parker. Parallel remote method invocation and m-by-n data redistribution. In *Proceedings of the 4th Los Alamos Computer Science Institute Symposium*, page (published on CD), 2003.

[20] S. Dillard, V. Natarajan, G. Weber, V. Pascucci, and B. Hamann. Topology-guided tessellation of quadratic elements. *International Journal of Computational Geometry and Applications (IJCGA)*. To appear.

[21] S. E. Dillard, V. Natarajan, G. H. Weber, V. Pascucci, and B. Hamann. Tesselletion of quadratic elements. In *ISAAC*, pages 722–731, 2006.

[22] C. Garth, F. Gerhardt, X. Tricoche, and H. Hagen. Efficient computation and visualization of coherent structures in fluid flow applications. *Transactions on Visualization and Computer Graphics, IEEE Visualization 2007*, 13(5), 2007.

[23] C. Garth, G.-S. Li, X. Tricoche, and C. D. Hansen. Interactive Visualization of Coherent Structures in Transient Flows. March 2007.

[24] A. G. Gyulassy, M. Duchaineau, V. Natarajan, V. Pascucci, E. M. Bringa, A. Higginbotham, and B. Hamann. Topologically Clean Distance Fields. *IEEE Transactions on Visualization and Computer Graphics*, 13(5):To Appear, 2007.

[25] L. M. Hwa, M. A. Duchaineau, and K. I. Joy. Adaptive 4-8 texture hierarchies. In *Proceeding of IEEE Visualization 2004*, pages 219–226, 2004.

[26] C. Johnson. Top scientific visualization research problems. *IEEE Computer Graphics and Applications: Visualization Viewpoints*, 24(4):13–17, July/August 2004.

[27] C. Johnson, M. Berzins, L. Zhukov, and R. Coffey. SCIRun: Applications to atmospheric diffusion using unstructured meshes. In M. Baines, editor, *Numerical Methods for Fluid Dynamics VI*, pages 111–122. Oxford University Press, 1998.

[28] C. Johnson and S. Parker. Applications in computational medicine using SCIRun: A computational steering programming environment. In H. Meuer, editor, *Supercomputer '95*, pages 2–19. Springer-Verlag, 1995.

[29] C. Johnson, S. Parker, and D. Weinstein. Large-scale computational science applications using the SCIRun problem solving environment. In *Proceedings of The International Supercomputer Conference 2000*, 2000.

[30] C. Jones, K.-L. Ma, A. Sanderson, and L. Myers. Visual Interrogation of Gyrokinetic Particle Simulations. 78, June 2007.

[31] K. I. Joy, M. Miller, H. Childs, E. W. Bethel, J. Clyne, G. Ostrouchov, and S. Ahern. Frameworks for Visualization at the Extreme Scale. 78, June 2007.

[32] J. Kniss, S. Premoze, M. Ikits, A. E. Lefohn, C. Hansen, and E. Praun. Gaussian transfer functions for multi-field volume visualization. In *IEEE Visualization 2003*, pages 497–504, October 2003.

[33] D. Laney, A. Mascarenhas, P. Miller, P. T. Bremer, and V. Pascucci. Understanding the Structure of the Turbulent Mixing Layer in Hydrodynamic Instabilities. *IEEE Transactions on Visualization and Computer Graphics*, 12(5):1053–1060, 2006.

[34] P. Lindstrom. Out-of-core simplification of large polygonal models. In K. Akeley, editor, *Proceedings of SIGGRAPH 2000*, Computer Graphics Proceedings, Annual Conference Series, pages 259–262, New Orleans, Louisiana, July 2000. ACM Press / ACM SIGGRAPH / Addison Wesley Longman.

[35] Luke Gosink and John C. Anderson and E. Wes Bethel and Kenneth I. Joy. Variable Interactions in Query-Driven Visualization. 13(6), October 2007.

[36] M. Miller, C. Hansen, and C. Johnson. Simulation steering with SCIRun in a distributed memory environment. In E. E. B. Kagstrom, J. Dongarra and J. Wasniewski, editors, *Applied Parallel Computing, 4th International Workshop, PARA'98*, volume 1541 of *Lecture Notes in Computer Science*, pages 366–376. Springer-Verlag, Berlin, 1998.

[37] M. Miller, C. Moulding, J. Dongarra, and C. Johnson. Grid-enabling problem solving environments: A case study of SCIRun and netsolv. In *Proceedings of The 5th International Conference on High-Performance Computing*, 2001 Advanced Simulation Technologies Conference, pages 98–103. Society for Modeling and Simulation International, April 2001.

[38] P. Miller, P.-T. Bremer, W. Cabot, A. Cook, D. Laney, A. Mascarenhas, and V. Pascucci. Application of morse theory to analysis of rayleigh-taylor topology. In *10th International Workshop on the Physics of Compressible Turbulent Mixing*, 2006.

[39] OMG. Corba Component Model, visited 1-11-2000. http://www.omg.org/cgi-bin/doc?orbos/97-06-12.

[40] A. Pang, C. Hansen, and M. Chen, editors. *IEEE Visualization Conference and IEEE Information Visualization Conference Proceedings 2007.* IEEE Transactions on Visualization and Computer Graphics. IEEE, November 2007.

[41] S. Parker, D. Beazley, and C. Johnson. Computational steering software systems and strategies. *IEEE Computational Science and Engineering*, 4(4):50–59, 1997.

[42] S. Parker and C. Johnson. SCIRun: A scientific programming environment for computational steering. In *Supercomputing '95*. IEEE Press, 1995.

[43] S. Parker, M. Miller, C. Hansen, C. Johnson, and P.-P. Sloan. An integrated problem solving environment: The SCIRun computational steering system. In H. El-Rewini, editor, *31st Hawaii International Conference on System Sciences (HICSS-31)*, volume VII, pages 147–156. IEEE Computer Society, January 1998.

[44] S. Parker, D. Weinstein, and C. Johnson. The SCIRun computational steering software system. In E. Arge, A. Bruaset, and H. Langtangen, editors, *Modern Software Tools in Scientific Computing*, pages 1–40. Birkhauser Press, Boston, 1997.

[45] S. Parker, K. Zhang, K. Damevski, and C. Johnson. *Integrating Component-Based Scientific Computing Software*, page (accepted). 2005.

[46] S. G. Parker, K. Damevski, A. Khan, A. Swaminathan, and C. R. Johnson. *Advanced Computational Infrastructures for Parallel/Distributed Adaptive Applications*, chapter The SCIJump Framework for Parallel and Distributed Scientific Computing. Wiley Press, 2007. to appear.

[47] V. Pascucci, G. Scorzelli, P.-T. Bremer, and A. Mascarenhas. Robust on-line computation of reeb graphs: simplicity and speed. *ACM Trans. Graph.*, 26(3):58, July 2007.

[48] A. Sanderson, M. Meyer, R. Kirby, and C. Johnson. A framework for exploring numerical solutions of advection reaction diffusion equations using a gpu based approach. *Journal of Computing and Visualization in Science*, 2007. to appear.

[49] C. Scheidegger, D. Koop, H. Vo, J. Freire, and C. Silva. Querying and Creating Visualizations by Analogy. *Transactions on Visualization and Computer Graphics, IEEE Visualization 2007*, 13(5), 2007.

[50] W. Schroeder, K. Martin, and W. Lorensen. *The Visualization Toolkit.* Prentice-Hall Inc., 1996.

[51] J. Shepherd and C. Johnson. Hexahedral mesh generation constraints. *Engineering with Computers*, 2007. to appear.

[52] K. Stockinger, J. Shalf, K. Wu, and E. W. Bethel. Query-driven visualization of large data sets. In *Proceedings of IEEE Visualization 2005*, pages 167–174, 2005.

[53] H. T. Vo, S. P. Callahan, P. Lindstrom, V. Pascucci, and C. T. Silva. Streaming simplification of tetrahedral meshes. *IEEE Transactions on Visualization and Computer Graphics*, 13(1):145–155, 2007.

[54] I. Wald, H. Friedrich, A. Knoll, and C. Hansen. Interactive Isosurface Ray Tracing of Time-Varying Tetrahedral Volumes. *Transactions on Visualization and Computer Graphics, IEEE Visualization 2007*, 13(5), 2007.

[55] G. Weber, S. Dillard, H. Carr, V. Pascucci, and B. Hamann. Topology-Controlled Volume Rendering. *IEEE Transactions on Visualization and Computer Graphics*, 13(2):330–341, March/April 2007.

[56] G. H. Weber, V. Becker, H. Childs, T. Ligocki, B. van Straalen, and E. W. Bethel. Visualization Tools for Adaptive Mesh Refinement Data. In W. Benger and R. Heinzel and W. Kapferer and W. Schoor and M. Tyagi and S. Venkataraman and G.H. Weber, editor, *Proceedings of the 4th High End Visualization Workshop*, pages 12–25, June 2007.

[57] G. H. Weber, P.-T. Bremer, and V. Pascucci. Topological Landscapes: A Terrain Metaphor for Scientific Data. *IEEE Transactions on Computer Graphics, Proceedings of Visualization 2007*, 2007.

[58] A. Wiebel, X. Tricoche, D. Schneider, H. Jänicke, and G. Scheuermann. Generalized Streak Lines: Analysis and VIsualization of Boundary Induced Vortices. *Transactions on Visualization and Computer Graphics, IEEE Visualization 2007*, 13(5), 2007.

[59] S.-E. Yoon and P. Lindstrom. Random-Accessible Compressed Triangle Meshes. *IEEE Transactions on Visualization and Computer Graphics*, 13(5):To Appear, 2007.

[60] K. Zhang, K. Damevski, V. Venkatachalapathy, and S. Parker. SCIRun2: A CCA framework for high performance computing. In *Proceedings of The 9th International Workshop on High-Level Parallel Programming Models and Supportive Environments*, April 2004.